BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF INFORMATION TECHNOLOGY
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF INFORMATION SYSTEMS
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

PROCESSING AND VISUALIZATION OF PSYCHOLOGICAL DIAGNOSTIC DATA

ZPRACOVÁNÍ A VIZUALIZACE PSYCHOLOGICKÝCH DIAGNOSTICKÝCH DAT

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR — **Bc. ANDREA CHIMENTI**
AUTOR PRÁCE

SUPERVISOR — **Ing. JIŘÍ HYNEK, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2024

Master's Thesis Assignment

153667

Institut: Department of Information Systems (DIFS)
Student: **Chimenti Andrea, Bc.**
Programme: Information Technology and Artificial Intelligence
Specialization: Machine Learning
Title: **Processing and Visualization of Psychological Diagnostic Data**
Category: Information Systems
Academic year: 2023/24

Assignment:

1. Study the issues of psychological testing and the existing types of psychological tests.
2. Study the principles of information systems development and data processing.
3. In cooperation with the Rehabilitation Institute Kladruby, analyze the requirements for processing and evaluating data from psychological testing.
4. According to the results of the analysis, design appropriate views for visualization of testing results (detailed views, comparison of patients according to different attributes, aggregated statistics).
5. Implement the proposed views using web technologies.
6. Test the usability of the implemented solution in cooperation with the Rehabilitation Institute Kladruby.

Literature:

- Ferjenčík, J. (2000). Úvod do psychologického výzkumu. Portál, Praha.
- Hogan, T. P. (2013). Psychological Testing: A Practical Introduction. Wiley.
- Kirk, A., et al. (2016). Data Visualization: Representing Information on Modern Web. Packt Publishing.
- Svoboda, M. (1999). Psychologická diagnostika dospělých. Portál, Praha.
- Tufte, E. R. (2001). The Visual Display of Quantitative Information. Graphics Press.
- Urbánek, T. (2011). Psychometrika: měření v psychologii. Vyd. 1. Praha: Portál, 2011. 319 s.
- Wilke, C. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures. O'Reilly Media.

Requirements for the semestral defence:
Items 1 - 4.

Detailed formal requirements can be found at https://www.fit.vut.cz/study/theses/

Supervisor: **Hynek Jiří, Ing., Ph.D.**
Consultant: Fiľová Petra, PhDr., Ph.D.
Head of Department: Kolář Dušan, doc. Dr. Ing.
Beginning of work: 1.11.2023
Submission deadline: 17.5.2024
Approval date: 30.10.2023

Abstract

This thesis presents an information system developed in collaboration with the Rehabilitation Center Kladruby. The system's objective is to digitalize and unify the process of collecting, evaluating, and visualizing psychological diagnostic data. The system is implemented as a web application with a three-tier architecture using FastAPI and Vue.js. The application allows psychologists to record testing results, convert between different types of scores, and create visualizations to compare patients. Additionally, the application incorporates a module for exporting data, facilitating further analysis or research projects. The application has successfully undergone user acceptance testing.

Abstrakt

Tato práce představuje informační systém vyvinutý ve spolupráci s Rehabilitačním ústavem Kladruby. Cílem systému je digitalizace a sjednocení procesu sběru, vyhodnocování a vizualizace psychodiagnostických dat. Systém je implementován jako webová aplikace s třívrstvou architekturou pomocí knihoven FastAPI a Vue.js. Aplikace umožňuje psychologům zaznamenávat výsledky testů, převádět mezi různými druhy skórů a vytvářet vizualizace pro porovnávání pacientů. Dále obsahuje modul pro export dat, který usnadňuje provádění další analýzy nebo výzkumných projektů. Aplikace úspěšně prošla uživatelským akceptačním testováním.

Keywords

psychometrics, psychological testing, test norms, Rehabilitation Center Kladruby, web application, information system, visualization, FastAPI, Vue.js

Klíčová slova

psychometrika, psychologické testování, testové normy, Rehabilitační ústav Kladruby, webová aplikace, informační systém, vizualizace, FastAPI, Vue.js

Reference

CHIMENTI, Andrea. *Processing and visualizationof psychological diagnostic data.* Brno, 2024. Master's thesis. Brno University of Technology, Faculty of Information Technology. Supervisor Ing. Jiří Hynek, Ph.D.

Rozšířený abstrakt

Tato práce se zabývá návrhem a implementací informačního systému, který umožní efektivní sběr psychologických diagnostických dat a jejich následné zpracování a vyhodnocování. Informační systém vznikl ve spolupráci s Rehabilitačním ústavem Kladruby, který patří mezi přední česká zdravotní střediska podobného druhu. Primární motivací ke vzniku tohoto informačního systému je zájem rehabilitačního ústavu o digitalizaci procesu zpracování a vyhodnocování dat psychologických vyšetření. Dále se jedná o příslib vzniku databáze mimořádně statisticky robustního vzorku údajů o specifikách úrovně kognitivních funkcí pacientů, který je mimořádný nejen v rámci České republiky, ale i Evropy. Tato databáze bude sloužit jako základ výzkumných projektů na univerzitách s psychologickým a zdravotnickým zaměřením.

Úvodní kapitola práce se zabývá testováním v psychologii, které hraje klíčovou roli ve stanovování diagnóz a v procesu vyhodnocování celkového zdravotního stavu jedince. Výsledky psychologických testů jsou zaznamenávány v podobě hrubých skórů, které obvykle samy o sobě nenesou žádný specifický význam. Proto, aby mohly být hrubé skóry korektně interpretovány, je nutné je porovnat s referenční skupinou, která má podobné parametry jako testovaný jedinec. Obor psychometrie za tímto účelem definuje tzv. normy. Kvalita použitých norem má přímý vliv na přesnost vyhodnocení provedeného testu. V případě, že je norma nepřesně stanovená nebo definovaná na vzorku populace, který neodpovídá testovanému jedinci, mohou být výsledné interpretace skórů zavádějící. Tento jev není výjimečný, jelikož mnoho norem vzniklo na malém vzorku populace. V některých případech mohou normy pro danou referenční skupinu i zcela chybět. Tato skutečnost zdůrazňuje důležitost sběru psychodiagnostických dat, která poskytují nezbytný základ pro vývoj psychometrických norem. Kapitola podrobně popisuje proces vyhodnocování testových výsledků pomocí norem a nezbytný matematický aparát který je nutný k jeho pochopení.

Druhá kapitola práce poskytuje teoretický základ tvorby informačních systémů se zaměřením na vizualizaci dat. Kapitola nejprve formálně definuje informační systém a poté popisuje známé architektury, přesněji dvou-vrstvou a tří-vrstvou architekturu. Dále jsou popsány rozdíly mezi daty, informacemi, znalostmi a moudrostí, které umožňují lépe pochopit roli dat v informačních systémech. Text se dále zabývá explorativní datovou analýzou a principy vizualizace dat. Důležitost vizualizace dat je znázorněna na Anscombeho kvartetu, tedy čtveřici grafů s totožnými statistickými charakteristikami, ale zcela odlišnými vizualizacemi. Principy správného návrhu vizualizace jsou popsány dvěma soubory pravidel a to Gestalt principy vizuálního vnímání a Tufteho principy grafické integrity. Konec kapitoly představuje nejpoužívanější typy grafů, které je možné použít pro vizualizaci dat.

Následující část práce již popisuje kroky, které vedly k samotnému vzniku systému. Prvním krokem bylo objasnění požadavků na výsledný systém. Tyto požadavky vznikly na základě konzultací s reprezentanty z Rehabilitačního ústavu Kladruby. Základním cílem je usnadnění procesu vyhodnocování psychologických testů, které by navíc umožňovalo data vizualizovat za účelem porovnání pacientů. Díky již existujícímu proof of concept prototypu systému, který vznikl v rámci předmětu Informační systémy na Fakultě informačních technologií Vysokého učení technické v Brně, bylo možné i technickou část požadavků definovat velmi přesně.

Na základě získaných požadavků byla navržena aplikace založena na tří-vrstvé architektuře, která umožňuje zcela oddělit klientskou aplikaci od serveru. Serverová část se skládá z relační databáze a REST API, které poskytuje komunikační rozhraní definované OpenAPI specifikací. Práce podrobně představuje seznam navržených koncových bodů. Na straně klienta se jedná o webovou aplikaci, která nevyžaduje instalaci přídavného softwaru a umožňuje tak rychlejší nasazení celkového řešení. Kapitola dále popisuje návrh vizualizací a datových matic, které slouží pro export dat.

Další kapitola rozebírá technické detaily implementace. Databáze využívá systému řízení dat PostgreSQL. REST API bylo implementováno pomocí knihovny FastAPI pro Python a skládá se ze tří vrstev, které jsou podrobněji popsány v textu práce. Klientská aplikace využívá frameworku Vue.js a UI knihovny Vuetify, díky které se uživatelské rozhraní řídí principy Material Design.

Finální část práce se zabývá procesem testování. Aplikace prošla akceptačním uživatelským testováním, které ověřilo že se jedná o použitelné řešení. Na základě tohoto testování bylo navrženo několik změn, jejichž většina byla do aplikace implementována.

Práce přináší plně funkční řešení, které je až na pár drobnějších nedodělků nasaditelné. Aplikace přinese zefektivnění procesu sběru a vyhodnocování psychodiagnostických dat a umožní zvýšit úroveň individuální péče v Rehabilitačním ústavu Kladruby. V případě dlouhodobé spokojenosti s aplikací a naplnění výzkumných cílů může být aplikace nasazena na zdravotnických ústavech s podobným zaměřením.

Processing and visualizationof psychological diagnostic data

Declaration

I hereby declare that this Masters's Thesis was prepared as an original work by the author under the supervision of Ing. Jiří Hynek, Ph.D. The supplementary information was provided by PhDr. Petra Fiľová, Ph.D. I have listed all the literary sources, publications and other sources, which were used during the preparation of this thesis.

. . . .
Andrea Chimenti
May 16, 2024

Acknowledgements

I want to express my gratitude to my supervisor, Ing. Jiří Hynek, Ph.D., for his professional guidance, patience, and flexibility. I also wish to thank my consultant, PhDr. Petra Fiľová, Ph.D., who provided me with the psychological background necessary for this thesis.

I would also like to thank my parents and family for giving me the opportunity to pursue and complete my studies. Special thanks go to those who supported me during the most challenging periods, particularly Aleš Chudárek, Natálie Hlavatá, Jan Klhůfek, and Hedvika Kotherová (in alphabetical order).

Contents

1 Introduction **3**

2 Testing in Psychology **4**
2.1 Tests and Their Types 4
2.2 Classical Test Theory 6
2.3 Underlying Distributions 8
2.4 Norms 11

3 Information Systems and Data Visualization **18**
3.1 Defining an Information System 18
3.2 Architectures of Information Systems 20
3.3 Data, Information, Knowledge and Wisdom 23
3.4 Exploratory Data Analysis 23
3.5 Data Visualization Principles 25
3.6 Chart Types 32

4 Analysis of Requirements **41**
4.1 Rehabilitation Center Kladruby 41
4.2 Workflow 42
4.3 Research Goals 43
4.4 Test Management 44
4.5 Test Result Analysis 46
4.6 System Requirements 47

5 Design **50**
5.1 Architecture 50
5.2 User Management 51
5.3 Data Model 52
5.4 REST API 53
5.5 Client 55
5.6 Data Visualizations 55
5.7 Exports 56

6 Implementation **59**
6.1 Backend 59
6.2 Frontend 60
6.3 Authentication and Authorization 62

7 Testing **64**
7.1 Unit Testing 64
7.2 User Acceptance Testing 65

8 Conclusion **67**

Bibliography **68**

A User Interface of the Final Application **71**

B Prototype **75**

C Tests Used at the Rehabilitation Center Kladruby **77**

D Visualization Design Protocols **78**

E REST API Endpoints **80**

Chapter 1

Introduction

Psychological testing plays a critical role in the evaluation process of an individual's medical state. The results of tests are collected in raw format, which typically has little intrinsic meaning. Frequently, questions such as "Does a raw score of 23 points indicate an above-average performance, is it merely average, or does it suggest underperformance?" arise. To interpret these scores meaningfully, they must be compared with a reference group. In the field of psychometrics, introduced in Chapter 2, these comparisons are made through norms. The quality of norms directly impacts the efficiency of interpretation. However, norms are often outdated, imprecise, or entirely absent, leading to the necessity of introducing new ones. For new norms to accurately reflect reality, they require a substantial amount of data. This underscores the importance of properly collecting psychological diagnostic data, which provides the necessary information for the development of such norms.

The Rehabilitation Center Kladruby is a medical facility that provides comprehensive care for patients with neurological and musculoskeletal disorders, post-surgical conditions, spinal injuries, and other medical conditions [30]. The diverse patient population at the center provides a robust dataset that is ideal for the creation of new norms. However, the center's current data management system is inadequate, leading to an inefficient use of data. Furthermore, the clinicians' ability to effectively evaluate and analyze psychological and neurological test results, a crucial step in delivering personalized care, is limited by the deficiencies of the existing data management system.

The objective of this thesis is to develop an information system in collaboration with the Rehabilitation Center Kladruby that will allow to effectively collect diagnostic data and enhance the manner in which these data are analyzed within the center. The anticipated outcome is an increase in the usability of test data at the Rehabilitation Center Kladruby and the potential creation of a robust database of diagnostic data that will be utilized for research purposes. Chapter 3 provides a theoretical foundation for the concepts of information systems and data visualizations. Chapter 4 then covers in detail the requirements that emerged for this system from discussions with representatives of the center.

The subsequent chapters describe the development process of the information system. Chapter 5 introduces the design of the application, focusing on the architectural layers, the data model, and the data visualization techniques employed. Chapter 6 presents the details of the implementation. Finally, Chapter 7 describes the testing process, which was conducted in two sessions with representatives from the Rehabilitation Center Kladruby.

Chapter 2

Testing in Psychology

This chapter introduces the fundamental principles of the clinical use of psychological tests. Testing in psychology is defined as the process of measuring and evaluating various aspects of human behavior using standardized tools in the form of questions and tasks that are administered to the tested subject. The aim of clinical testing is to assist an individual who may have or has some type of problem. Testing helps to identify the nature and severity of the problem and also helps to monitor the progress [19]. A variety of tests exist for this purpose, some of which are discussed in Section 2.1.

The field that deals with the issues of testing in psychology is called psychometrics. It is an inseparable part of psychodiagnostics and plays a crucial role in the process of creating and evaluating psychological tests [38]. Psychometrics covers both the theoretical and practical aspects of testing and ultimately defines conceptual frameworks that are the applied to the process of testing. One of the foundational frameworks, Classical Test Theory, is discussed in Section 2.2.

The evaluation of psychological tests relies heavily on mathematical and statistical methods. Clinicians use reference groups to contextualize the results of a patient. It is essential for them to understand the laws that influence the properties of the results of the reference groups. Section 2.3 introduces some of the most commonly encountered probability distributions in the field of psychology. The interpretation of test results is based on norms. Norms are an essential part of the evaluation process as they dictate how to correctly interpret the results of patients. Section 2.4 describes the usage of norms and introduces some of the most common types.

2.1 Tests and Their Types

In 1949, Lee Joseph Cronbach, an influential American educational psychologist, defined a psychological test as follows [14]:

> *A test may be defined as a systematic procedure for comparing the behavior of two or more persons.*

This definition, which is broad in scope, remains relevant in the field to this day. As [16] states, the definition includes three important components. First, tests involve behavioral samples of some kind. Second, the behavioral samples must be collected in a systematic and and standardized manner. Finally, tests are designed to identify differences between individuals, including comparisons of performance of a single individual over time.

In clinical practice, psychological tests are used as a means of assessing various psychological attributes, such as intelligence, depression, aptitude, extroversion, and many more. Since these attributes are not directly observable, psychologists must identify some type of observable behavior that they believe represents the particular unobservable psychological attribute, state, or process. They then use tests to sample the behavior and attempt to interpret those measurements in terms of the latent psychological characteristics that they think are reflected in the behavior [16].

Tests can be classified into various categories according to the measured psychological attributes. However, this classification is not rigid, as tested characteristics often have some overlap with each other. The following categorization is adapted from [19].

2.1.1 Mental Ability Tests

Mental ability tests are designed to assess a person's general cognitive ability or potential. The term *mental ability* encompasses a wide range of cognitive functions, including memory, spatial visualization, numerical ability, abstract reasoning, creative thinking, quantitative reasoning, and more. These tests typically measure traits that are considered to be signs of general intelligence. Classical examples of individually administered intelligence tests are the *Wechsler Adult Intelligence Scale* (WAIS) and the *Stanford-Binet Intelligence Scale*. Mental ability tests can also be administered to entire groups, such as school classes. One example of such a test is the Scholastic Aptitude Test, which is used in the United States to predict success in college [19].

2.1.2 Achievement Tests

Achievement tests are used to determine a subject's level of knowledge or skill in a specific domain. These tests may cover broad topics, such as mathematics, reading, writing, social studies, or more specific topics, such as Italian, genetics, or calculus. Certification or licensing examinations are additional subcategories of achievement tests. If we consider tests that are not professionally assessed, this is arguably the most common type, since most school tests could fall under this category [19].

2.1.3 Personality Tests

This category comprises various tests designed to yield information about the human personality. They can be divided into two subcategories: objective personality tests and projective techniques. Objective personality tests attempt to measure objective personality traits using extended sets of closed-ended questions. The traits measured may be either within the normal range, such as extroversion, agreeableness, openness, etc., or within the pathological range, such as paranoia, hysteria, hypochondria. The responses are then compared to those of other clinical groups, and the final score is commonly expressed as a value on a scale [19].

Projective techniques use unstructured tasks and ambiguous stimuli to reveal the examinee's traits from their solution. These methods aim to evoke responses that may reveal aspects of the subject's personality by projecting internal attitudes, traits, and behavior patterns onto the external stimuli. The most famous example is the *Rorschach Inkblot Test*, where the tested subject uses their imagination to interpret the meaning of seemingly nonsensical blobs. Other examples are sentence completion techniques or human figure drawings. However, there is considerable controversy regarding the usefulness and reliability of projective tests. Psychologists hold divergent views on their value, as the results are frequently influenced by the subjective perception of the examiner interpreting the responses [11].

2.1.4 Neuropsychological Tests

These tests are designed to measure the functioning of the brain and the central nervous system (CNS) in general. They assist neuropsychologists in identifying the underlying cause of patients' problems. Tests in this category frequently overlap with other categories, particularly mental ability tests, as the health of the CNS directly affects cognitive functions such as memory, coordination, cognition, etc. From this perspective, this should not be considered a separate category, as many of the tests used in neuropsychological testing originate from other categories. However, a separate category is employed to encompass a set of tests that specifically measure brain functions [19]. The following list enumerates the subcategories of neuropsychological tests and provides examples of the tasks conducted within each subcategory [41]:

- **Memory tests:** Repetition of words, phrases, or numbers.
- **Cognition tests:** Explanation of relationships between two objects or abstract concepts.
- **Verbal communication tests:** Naming objects, listing words that begin with a given letter.
- **Motor tests:** Insertion of pegs into a pegboard using only one hand.

2.2 Classical Test Theory

Classical Test Theory (CTT) is a framework of concepts and techniques that has been used to develop various measurement instruments. Health researchers often deal with phenomena that are not directly observable, such as psychological traits or the extent of brain damage. To estimate these phenomena, they use indirect methods that rely on observable data, but these methods inevitably produce measurement error. One of the first theorists to acknowledge this problem was Charles Spearman, who clarified the relation between the observed data and the unobservable variable of interest. His work established the foundations of CTT [15].

The core idea of CTT is that an observed score is composed of the true state of the unobservable variable of interest and of an error caused by all other influences on the observable variable. The true state of the unobserved variable is it hypothetical and it's given by the equation [15]:

$$X = T + E \tag{2.1}$$

where:

$\boldsymbol{X}$ is the observed score.

$\boldsymbol{T}$ s the true value of the variable.

$\boldsymbol{E}$ is the observation error.

The observed variable is thus a combination of relevant information and error. Errors are assumed to be random and independent of true scores and other errors. Errors are also assumed to have a mean of zero and a normal distribution. Thanks to this assumptions, CTT provides methods for estimating and correcting for errors in order to improve the quality and usefulness of test scores and interpretations [15].

It should be noted that CTT is not the only existing conceptual framework, but it is certainly the most commonly encountered in test manuals, professional journals, and test score reports. Another popular framework is the Item Reponse Theory (IRT), for which a detailed coverage is beyond the scope of this work. In very simplified terms, the main difference between CTT and IRT lies in the assumption of CTT that all test items contribute equally to the test-taker's performance, which IRT does not [19]. CTT defines two key concepts that are used to evaluate the quality and usefulness of test scores and interpretations: reliability and validity. These concepts are often considered as the two main criteria for judging the adequacy of a psychological test for a certain purpose. It is not uncommon for these terms to be confused. However, it is crucial to differentiate between them. Subsections 2.2.1 and 2.2.2 cover their meaning in greater detail.

2.2.1 Validity

Validity refers to the extent to which a test measures what it claims to measure. More precisely it is the correctness of making assumptions based on the score obtained in the test. The crucial aspect is the purpose of the test. If a test is designed to measure depression, than the result should be a valid and reliable indicator of the patient's level of depression. It is important to note that validity is not an all-or-nothing concept, but a matter of degree. Determining the degree of validity of a test is a challenging task in psychology. Although counter-intuitive, it is possible for a test to be valid despite the lack of precision in its norms [19].

2.2.2 Reliability

Reliability refers to the stability and consistency of test scores. It indicates whether the score would be the same on different days. A reliable test consistently yields the same or similar score for an individual across repeated administrations. It is important to differentiate between a real change in the trait being measured and a random change that could depend on the actual state of the subject and the questions asked. The boundary between short-term and long-term changes is blurry and it's up to the examiner to decide the impact of the change in score on the tested subject's real traits. A real change in trait does not affect test reliability.

Reliability deals only with unsystematic errors in measurements and not constant errors. Constant errors are a type of error that causes the subject to consistently score with a large deviation. An extreme case would be taking an intelligence test in a foreign language that the subject does not know well. The score of the subject would be always low and the result would not reflect the true intelligence. It is also important to note that tests may be reliable without being valid, meaning that they may measure something consistently but not what they are supposed to measure [19].

2.3 Underlying Distributions

Understanding the underlying distribution of the results across a population of similar individuals is a key component in the process of creating and evaluating psychological tests. One of the most important distributions in statistics is the normal distribution [38], as many natural phenomena tend to follow this distribution. For example, given a population of people with the same sex and age, attributes such as short term memory capacity, reaction time, analytical thinking and many others tend to be normally distributed. This tendency toward normality in nature is so ubiquitous that when the underlying distribution of a variable is unknown, it is reasonable to assume that it is probably normal or at least close to it. Just like these natural phenomena, psychological tests and surveys also tend to approximate the normal distribution [19]. However, test scores can deviate from normality. In such case, it is sometimes possible to investigate and disclose the reason of the deviation. The true underlying distribution may follow one of the common cases introduced and discussed in this section[1].

2.3.1 Normal Distribution

The normal distribution, displayed in Figure 2.1, exhibits a number of important properties. First, it is characterized by its mean and standard deviation, which are incorporated into the formula presented in Equation 2.2 [18]. It is bell-shaped and symmetrical, with the left half being a mirror image of the right half. The mean, median, and mode are equal and are located at the center of the distribution. The distribution is unimodal, meaning that only one local mode exists and is equal with the global mode. The percentage of scores between the mean and a multiple of the standard deviation is equal for all normal curves [21]. This principle allows for the estimation of the number of individuals in each interval [38] and is shown in Figure 2.1.

$$\mathcal{N}(x;\mu,\sigma^2) = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \tag{2.2}$$

where:

$\boldsymbol{\mathcal{N}(x;\mu,\sigma^2)}$ is the probability distribution function of the normal distribution.

$\boldsymbol{\mu}$ is the mean of the distribution.

$\boldsymbol{\sigma^2}$ is the variance of the distribution, defined as the second power of the standard deviation.

Although several methods are based on the assumption that the distributions they work with are normal, it is important to note that a perfectly normal distribution is a theoretical ideal. In reality, true distributions of actual test scores are not perfectly normally distributed. For example, in the distribution of real IQ scores, there may be a few more people who are on either the lower or upper side of the scale. Such distributions are skewed, but the degree of skewness is often negligible, allowing for the assumption that the underlying distribution is normal [16]. However, there are instances when the deviation is significant enough to require a more detailed investigation of its cause.

[1]This section assumes that the reader is familiar with the following terms from descriptive statistic: mean, median, mode and standard deviation.

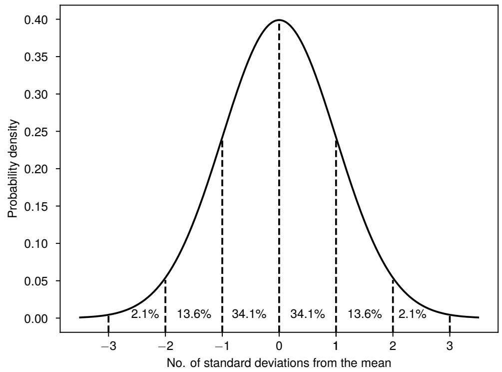
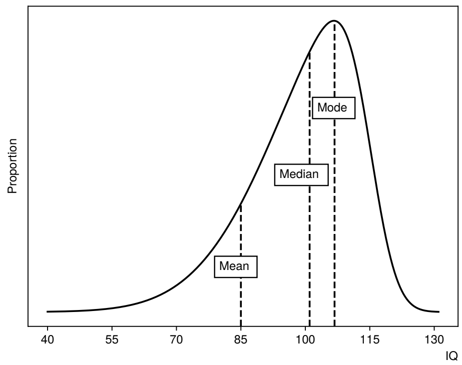


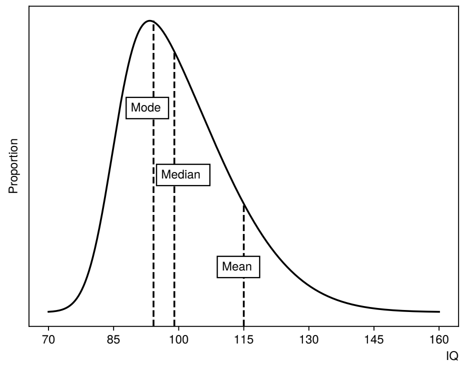
Figure 2.1: This illustration demonstrates the most significant properties of the normal distribution. The percentages represent the proportion of data within each interval. Approximately 68.2% of all values are within one standard deviation of the mean, 95.4% within two standard deviations, and 99.6% within three standard deviations [38].

2.3.2 Skewed Distribution

In a skewed distribution, scores tend to cluster at one end or the other of the x-axis, with the tail of the distribution extending in the opposite direction. In a positively skewed distribution, a few individuals have extremely high scores that pull the distribution in the positive direction. This results in a substantial inflation of the mean, accompanied by a decrease in the reliability of this indicator. The median is also pulled toward the positive end, but significantly less than the mean. A notable difference between these two indicators is a reliable sign of skewness. Conversely, a negatively skewed distribution is pulled in the negative direction by a set of individuals who have extremely low scores [21]. The difference between a negatively and a positively skewed distribution is illustrated in Figure 2.2.



(a) A negatively skewed distribution. The mean is pulled toward the left by a few extremely low scores [21].



(b) A positively skewed distribution. The mean is pulled toward the right by a few extremely high scores [21].

Figure 2.2: Distributions with negative and positive skewness. The scale on the x-axis is provided for illustrative purposes only.

The simplest explanation for the occurrence of a skewed distribution is that the method used for testing is either too easy or too difficult. If the questions are inappropriately simple, most individuals will tend to answer correctly and score a higher result on the scale, thus shifting the mode toward the right and causing a negative skew. Conversely, if the answers are too difficult, the opposite will occur. If the population was sampled correctly and the results are skewed, it is necessary to revise the testing method [38].

2.3.3 Multimodal Distribution

A typical indicator of a multimodal distribution is the presence of more than one local mode as depicted in Figure 2.3. A special case of the multimodal distribution is the bimodal distribution, which is characterized by the presence of specifically two local modes. The occurrence of a multimodal distribution of results provides strong evidence that the population is not homogeneous in relation to the psychological attribute being measured or other attributes that may have an impact on the results. A bimodal distribution suggests that the population can be divided into two distinct groups. Each of these groups follows a normal distribution of scores, but they differ in mean and variability [38].

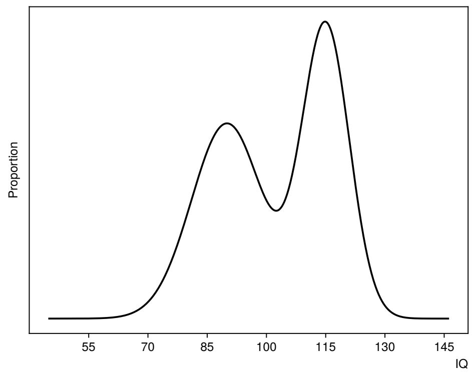


Figure 2.3: A bimodal distribution with two local maxima. The parameters of the distributions are μ=90, σ=9 for the first one, and μ=115, σ=6 for the second one. The scale on the x-axis is provided for illustrative purposes only.

2.3.4 L-shaped Distribution

The final type of distributions that can be frequently encountered in psychological testing, according to [38], are L-shaped distributions. They typically have a large number of occurrences (or data points) at the lower end of the range and very few at the higher end (resembling the shape of the letter L). An example of such a distribution can be seen in Figure 2.4. This type of distributions is rarely discussed as it is not exactly mathematically defined, but should be treated with caution. L-shaped distributions are not uncommon and are likely to be more prevalent than is generally recognized. They can significantly affect the performance of normality-assuming statistical tests and strongly resist transformation to normality. The thinner the tail of the distribution, the less likely it is that its L-shapedness will be detected by inspecting a sample drawn from it. Given the absence of a universally applicable rule of thumb to anticipate their occurrence, it is quite probable that a significant number of such distributions have gone undetected, with their data being improperly analyzed by normality-assuming statistics [10].

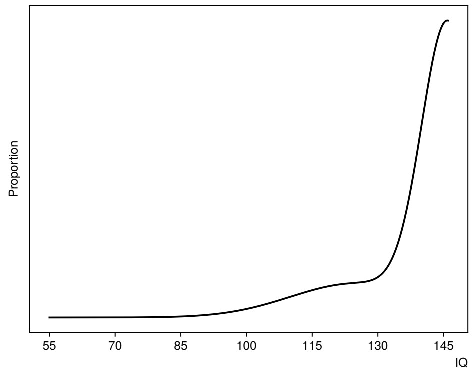


Figure 2.4: An example of a distribution that resembles the *L* shape. If this plot represented the distribution of the results of an intelligence test, it would be clear that the test is poorly designed and fails to unveil the real distribution of the attribute in the tested population. The scale on the x-axis is provided for illustrative purposes only.

One potential explanation for the occurrence of an L-shaped distribution in test results is a link between the tested attribute and social conformity. The majority of individuals comply with social norms and will select the same response. The less socially compliant the response is, the fewer individuals will select it. This phenomenon is also observed in the majority of medical disorders [38].

2.4 Norms

The interpretation of the scores obtained from a psychological test depends on relevant reference values. These reference values are called norms and they are based on the score of a representative sample of the population of interest. Subjects within a norm group demonstrate similar characteristics, like gender, age and ethnicity, among others [38]. The scores obtained from a test are called raw scores. They can represent different types of data such as time taken, number of tasks solved, percentage of questions answered with yes, etc. However, raw scores are not meaningful by themselves because they depend on the context. For instance, a score of 15 out of 20 correctly answered algebra questions may be considered normal for an adult, but extraordinary for an 8-year-old child. In order to make sense of raw scores, they must be converted into *normed* scores, which are also known as *derived* scores or *scale* scores. Normed scores are defined by comparing the raw score with the distribution of the raw scores of the norm group. Normed scores enable inferences about the level or degree of the attribute measured by the test [19].

The process of creating new norms can be understood as a search for the most suitable transformation of raw scores to normed. The more well-defined the transformation, the more accurate the interpretation of the results will be. The process of creating norms raises two key questions [38]. First, how should the reference population be defined? It is necessary to identify the factors that most precisely define the reference group. These include age, sex, socioeconomic status, region, occupation etc. There are two main methods for selecting a norm group. First, random sampling involves choosing subjects from a population with equal probability. The only requirement is defining the population clearly. A variation of this method is stratified random sampling, which divides the heterogeneous population into homogeneous subgroups based on relevant characteristics and then randomly samples from each subgroup. This method ensures that every subgroup is represented in the norm group and improves the accuracy of the estimates for the entire population [38]. In contrast, non-probabilistic sampling involves the selection of subjects based on deterministic criteria such as quotas, professional judgment, or the availability of volunteers. This method may be more convenient or practical, but it may also introduce biases or errors in the norm group [38].

The second question is how to compare the results of an individual with the results of the reference group [38]. This is defined by the type of the associated norm. Most of the norms commonly used in psychological tests fall into three broad types, each with its own possible subcategories. For many tests, several types of norms are available. Most of these types are related and can be converted from one type to another, although this is not a rule. Norms are typically conceptualized within the context of the normal distribution curve and experienced psychologists are able to swiftly make rough comparisons between test scores based on different types of norms [19]. Each of the three categories is discussed in greater detail in the following subsections.

2.4.1 Standard Scores

This group of derived scores is the most commonly used in psychological testing. Standard scores are based on z-scores, which express the value of the score as the distance from the mean in terms of multiples of the standard deviation. Z-scores and subsequently standard scores allow psychologists to directly compare results between two different tests, given that the tests gauge the same attribute and that the reference group used to convert the score is defined by the same parameters [38]. In order to convert the raw score into standard score, parameters of two normal distributions must be defined. The first distribution represents the distribution of raw scores of the reference group, and its parameters will be referred to as μ_r and σ_r in the following text. The second distribution, referred to as μ_s and σ_s, represents the distribution of standard scores in the desired resulting scale. The raw score is first converted into z-score. The subsequent step is to multiply the z-score by the standard deviation and add the mean of the scale. A graphical illustration of these steps can be found in Figure 2.5 and the complete formula is presented by the Equation 2.3 [19]. It is important to note that the amount of possible standard score systems is infinite, as the values of the mean and standard deviation can be chosen arbitrarily. However, the majority of applications of standard scores are based on a selection of the most well-known standard score systems. Some of these are presented in Table 2.1.

$$x_s = \sigma_s \cdot \underbrace{\frac{x_r - \mu_r}{\sigma_r}}_{z-score} + \mu_s \tag{2.3}$$

where:

- $\boldsymbol{x_s}$ is the standard score of the patient.
- $\boldsymbol{x_r}$ is the raw score of the patient.
- $\boldsymbol{\mu_s}$ is the mean in the standard score system.

$\boldsymbol{\mu_r}$ is the mean in the raw score system.

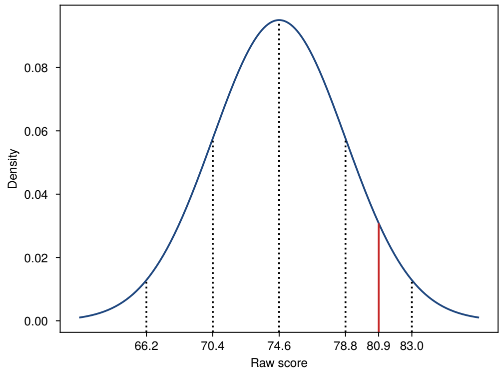
$\boldsymbol{\sigma_s}$ is the standard deviation in the standard score system.

$\boldsymbol{\sigma_r}$ is the standard deviation in the raw score system.

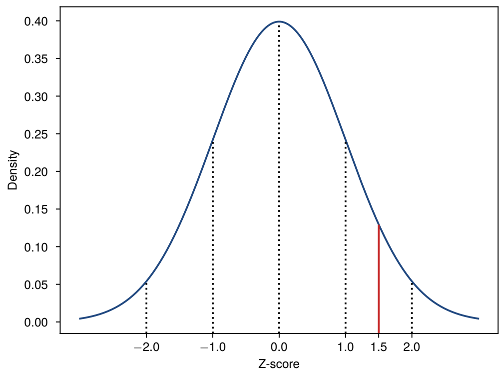
Score System	**Mean**	**Std. Dev.**	**Note**
T-score	50	10	T-scores are one of the most widely used. Thy have an effective range from 20 to 80 and are typically associated with personality tests [19].
Deviation IQs	100	15 (or 16)	Not to be confused with ratio IQ, deviations IQs are standard scores with an effective range from 55 to 145. They are typically associated with intelligence tests and they are well known to the public [19].
Stanine	5	2	Stanines are designed to divide the normal distribution into 9 units. They yield results in the range from 1 to 9 [38].
Sten	5.5	2	Same as stanines but they use 10 units instead of 9 [38].
Wechsler subtests	10	3	These scores are used in the subtests of the Wechsler Adult Intelligence Scale. The effective scale is from 1 to 19 [34].
CEEB	500	100	CEEB scores are used in Scholastic Aptitude Test for college admissions in the United States. The values of μ and σ are deliberately selected to be relatively high, allowing for a broader range of scores without the use of decimals [34].

Table 2.1: A brief overview of some of the popular standard score systems.

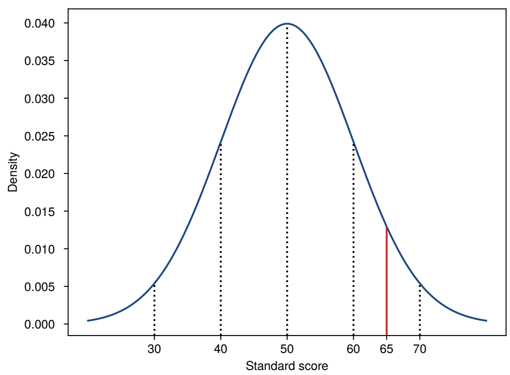
If the raw scores approximate a normal distribution, the conversion to standard scores is straightforward and the formula 2.3 is sufficient for this purpose. In the opposite case, however, the conversion becomes problematic. If the underlying distribution is at least assumed to follow a normal distribution, and the population is representative enough in terms of size, it is possible to use a method called *area transformation*. The core idea of this method is to divide the raw scores into intervals and to compute their percentiles. The raw scores are then mapped to a normal distribution based on the values of the percentiles [38].



(a) The raw score of the patient is compared with the distribution of raw scores of the reference population.



(b) The raw score is converted to a z-score to express its distance from the mean in terms of multiples of the standard deviation.



(c) The z-score is transformed using the selected standard scale. In this case the raw score is converted to a T-score.

Figure 2.5: An illustration of the standardization process, where the the z-score is used as a intermediary step [19].

2.4.2 Percentile Ranks and Percentiles

Another method for representing the results in psychometrics are percentiles and percentile ranks. The two terms are often used interchangeably, but they have a slightly different meaning. The percentile rank indicates the percentage of cases in the reference group that fall below a given raw score. This value is calculated based on the raw score and is expressed in percents. For instance, if a patient scores 15 points on a test and is assigned a percentile rank of 60, it means that 60% of other patients scored below a raw score of 15. On the other hand, a percentile represents a point on a scale below which a specified percentage of cases falls. The percentile is calculated based on the percentage and is expressed as a raw score value. Returning to the previous example, the 60th percentile of the test would be the raw score of 15 [19]. Percentiles and standard scores can be converted back and forth. The conversion is shown in Table 2.2.

Percentile	Z-score	IQ	T-score	CEEBS	Stanine
2	-2.0	70	30	300	1
7	-1.5	77.5	35	350	2
16	-1.0	85	40	400	3
31	-0.5	92.5	45	450	4
50	0.0	100	50	500	5
69	0.5	107.5	55	550	6
84	1.0	115	60	600	7
93	1.5	122.5	65	650	8
98	2.0	130	70	700	9

Table 2.2: Percentile equivalents of several standard score systems. The table contains only a subset of the most significant conversions. The complete table can be found in [19].

One of the most significant advantages of percentile ranks is their ease of comprehension. The majority of the population is able to interpret these scores correctly, as percentages are frequently used in western culture. For this reason, percentile ranks are the most commonly utilized method of communicating a score to patients in the field of psychology. However, percentile ranks also have a significant disadvantage that makes them vulnerable to misinterpretation. Since percentile ranks are distributed uniformly and test scores are typically distributed in a normal manner, the units at various points on the scale are not equal. Percentile ranks strongly reflect small changes in the raw score in the area close to the mean and weakly reflect substantial changes in the raw scare in both tails of the area [38]. This inequality is depicted in Figure 2.6.

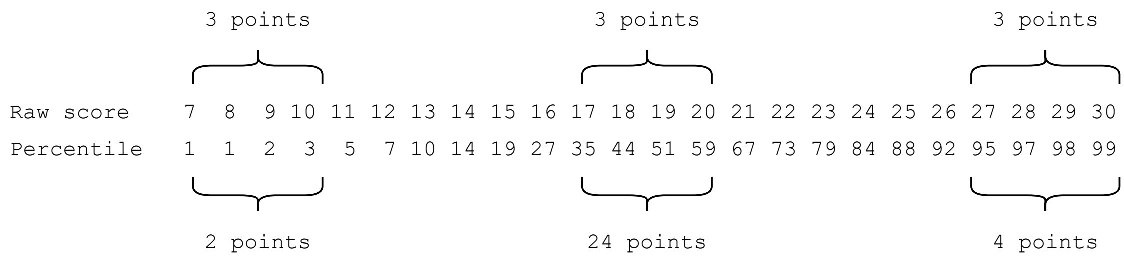


Figure 2.6: The inequality between the normal distribution of raw scores and the uniform distribution of percentile ranks [19].

2.4.3 Mental Age

The earliest types of norms used with psychological tests were mental ages. These were first developed and applied in 1908 (with a revision in 1911) by Alfred Binet and Théodore Simon in their intelligence test, the *Binet–Simon Scale*. The test was intended to measure the mental age of children and identify possible intellectual disabilities. The test defined a set of questions for each age group. The underlying assumption was that a child of a certain age would be able to complete the tasks assigned to their age and below. In the event that the child had performed poorly on tasks designed for their age, their true mental age would have been assumed to be lower than their chronological age [20].

Mental age scores are determined by comparing an individual's test performance to the average performance of individuals in the same age group. The average performances are plotted on a distribution curve, as seen in Figure 2.7. In practice, psychologists use conversion tables to convert a raw score into a mental age. The traditional definition of IQ (intelligence quotient), known as *ratio IQ*, is given by the formula in Equation 2.4 as the ratio of mental age to chronological age. Ratio IQs are used specifically with children, as they are less applicable for adults due to the non-linear nature of intellectual development. The concept of *ratio IQ* should not be confused with the previously discussed *deviation IQ*. Ratio IQs have become obsolete because it was observed that the standard deviations across age levels are not consistent, leading to misleading interpretations of the score [19].

$$IQ = \frac{A_m}{A_c} \cdot 100 \tag{2.4}$$

where:

$\boldsymbol{IQ}$ is the ratio IQ score.

$\boldsymbol{A_m}$ is the measured mental age.

$\boldsymbol{A_c}$ is the real chronological age.

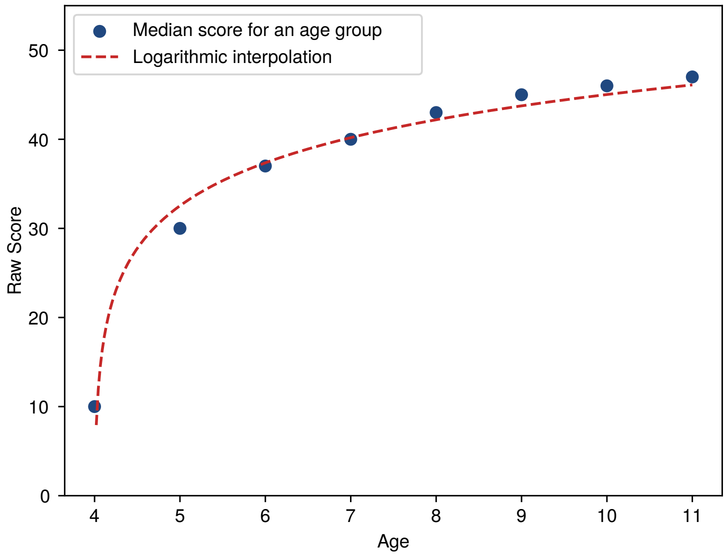


Figure 2.7: An example of the relationship between the median of the raw scores and the age of the tested subject [19]. The mental age linked to the raw score is found by feeding the raw score into the inversion of the interpolation function.

Chapter 3

Information Systems and Data Visualization

The following chapter provides a theoretical framework for understanding information systems and their capability to manage, process, and visualize data. Section 3.1 outlines the concept of an information system, its objectives and the activities encompasses. The following section, Section 3.2, introduces two of the most common architectures used to build information systems: the two-tier and three-tier architectures. Section 3.3, serving as a conceptual bridge between information systems and data processing, describes the hierarchy of the *Data, Information, Knowledge, Wisdom pyramid* and clarifies the role of data within information systems.

Section 3.4 examines the data-driven process of extracting information from data, a method particularly useful in environments with a substantial amount of pre-collected data. The principles of data visualization methods are discussed in Section 3.5. First, the motivational use case of Anscombe's Quartet is presented. Then, two different sets of principles are introduced: The Gestalt principles of visual perception and Tufte's Principles of Graphical Integrity. These principles form a robust foundation for data visualization. The last section, Section 3.6, presents specific visualization techniques that are commonly used by data analysts in their daily practice.

3.1 Defining an Information System

Giving an exact definition of an information system is a challenging task, due to varying opinions from different sources. However, a widely accepted definition of an information system can be found in [23]:

> *An information system can be defined technically as a set of interrelated components that collect (or retrieve), process, store, and distribute information to support decision making and control in an organization.*

A similar definition is found in [8]:

> *An information system is a system for collecting, processing, storing, retrieving, and distributing information within the enterprise and between the enterprise and its environment.*

However, the following definitions offer an alternative perspective on the topic of information systems.

> *An information system can be any organized combination of people, hardware, software, communications networks, data resources, and policies and procedures that stores, retrieves, transforms, and disseminates information in an organization.* [24]

> *Information systems are systems for using signs*[1] *in the sense that they act as a communication medium between different people, sometimes spatially and temporally distant.* [9]

A common pattern that is observable throughout all the definitions above and also coherent with [8] is that *"The primary objective of information systems is to provide and maintain an integrated flow throughout the organization or enterprise, so that the right information is available whenever a wherever needed, in the quality and quantity needed."* Information systems support decision making and coordination. They help workers analyze problems, visualize complex subjects or create new products [23].

According to [23], an information system can be described as a set of activities, which produce information required by the company. These activities include input, processing, and output. Input involves the collection of raw data from the external environment. Processing involves transformation processes that convert this raw data into a more useful form, such as information. Output transfers the produced information to end users. [23] also defines one last required component, feedback, which is the output that is returned to users in order to help them evaluate and correct the input stage.

That differs from the definition of feedback in [24], which states that feedback is *"data about the performance of a system*[2]*."* In addition to activities defined above, [24] also includes two more elements, control and storage. Control is defined as the monitoring and evaluation of feedback in order to determine whether a system is moving toward achieving its goal. In other words, it is the act of responding to feedback from the information system in the real world. Storage is an activity in which data are retained in an organized manner for later use. This organization facilitates their later use in processing or retrieval as output when need by the end user. The relationships between these activities are shown in Figure 3.1.

[1] In this context a sign is anything that is significant. Signs are the core element serving to link issues of human interactions, meaning, the structure of language, forms of communication transmission, data storage and collaborative action [9].

[2] In this context, a system is the real-world environment that information systems aim to model.

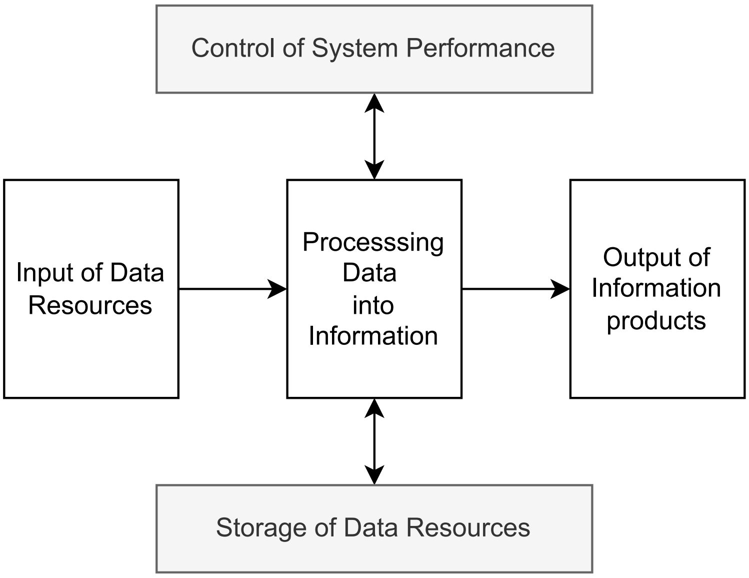


Figure 3.1: A diagram of control activities that process and transform data resources into information products [24].

3.2 Architectures of Information Systems

[8] defines the architecture of an information system as the *„integral structural design of a system, its elements and their relationships depending on given system requirements.“* Essentially understanding architecture as an abstract plan that includes the corresponding designing process of the system's structure appropriate to the goals of the system. The following subsections list and describe some of the most common architectures used in real world information systems. [27] is the primary source for the subsequent subsections on two-tier and three-tier architectures.

3.2.1 Two-Tier Client/Server

The aim of the two-tier client/server architecture is to enable a number of users to access the same data from different locations. The application is divided into two layers: the presentation layer, also known as client, and the data storage layer, also known as server. The data storage layer is centralized and and allows multiple workstations to have simultaneous access to data, which is usually stored on a database server. The setup is illustrated in Figure 3.2. In contrast, the presentation layer provides users with an interface on their workstations.

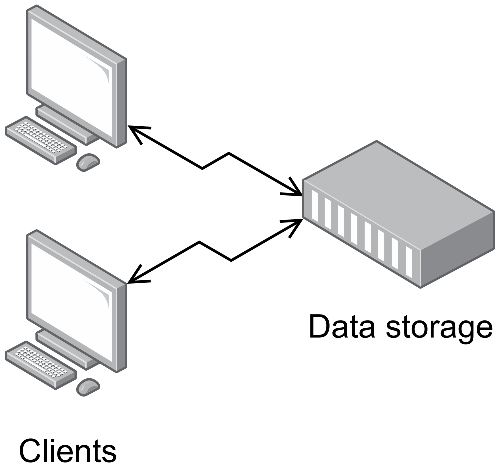


Figure 3.2: In the two-tier architecture, clients have direct access to the centralized data storage [27]. The data storage server is the only component responsible for handling client requests.

A simple application can be developed quickly using the two-tier architecture. Any database application that handles data retrieval and data presentation elsewhere can be considered a client/server application. The two-tier architecture is a straightforward solution for basic applications that require limited data processing and validation.

However, the two-tier architecture has a number of major limitations, particularly when greater complexity is required. The client communicates directly with the data store and is responsible for data processing. This results in several disadvantages:

- Two-tier systems are notorious for the inability to adapt to changing environments and scale with growing user and data volume. The client is loaded with knowledge completely unrelated to the task of data presentation. A change to the data model of the server requires all clients to update their pre-processing steps. This involves the modification, testing and redistribution of the client whose core functionality is supposed to be the presentation of the data. This leads to a rigid and inflexible application update process.
- As the application becomes more complex, additional needs may arise without a clear place to fit them. This can cause significant scope creep, which leads to unnecessarily heavy and fat clients.
- Since each client has its own copy of the data, if one client makes changes to the records on the server, other clients may end up with outdated data, resulting in a problem called “dirty writes”. Dirty writes occurs when a user tries to modify a record based on an old, out-of-date copy of the database.

To overcome the limitations of the two-tier architecture, a three-tier system was developed. The properties of this system are discussed in the following subsection.

3.2.2 Three-Tier

The three-tier architecture, sometimes called monolithic [32], introduces an additional layer, which centralizes data processing, to the two-tier model. This layer contains the business rules for processing data and isolates this process in a central location. The mid-tier layer is sometimes referred to as the application layer or the business layer. In this case naming conventions are not strict and may vary between sources. A schema of the model is shown in Figure 3.3.

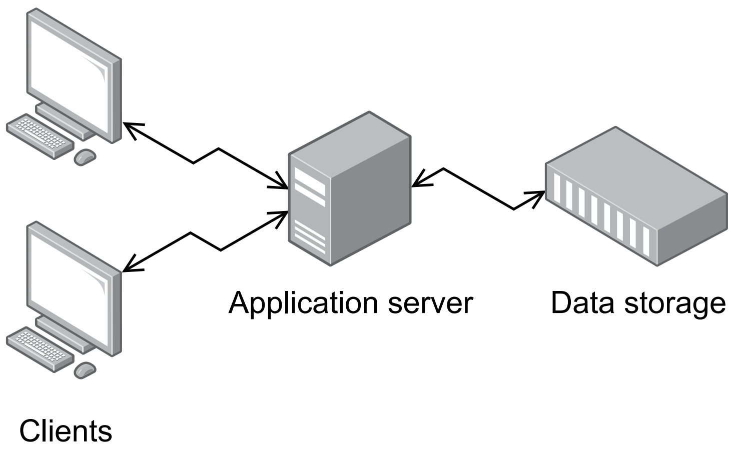


Figure 3.3: In the three-tier architecture, the application server acts as a connector between the clients and the data storage server, isolating them from each other. The figure is an adaptation derived from [27].

The application layer is responsible for handling the operations that do not fit into either of the two tiers of the two-tier architecture. It is usually implemented in the form of an application server that exposes a public API interface. The connection between the application layer and the database is isolated, and the presentation layer is unconcerned with the way data are stored. The client only has access to the entities in the application tier. The advantages of the three-tier architecture are summarized in the following list:

- It resolves the limitations of the two-tier architecture that were mentioned in the previous section.
- The client is independent on the data model and its scope is reduced to data presentation.
- The business layer improves the security of the application by controlling the clients' access to resources.
- The business rules are centralized in one place, making them easier to manage.

The three-tier architecture is a natural evolution from the two-tier model and provides considerable advantages. However, it is not a universally applicable solution, particularly in the context of modern cloud systems. Some sources, such as [32], go as far as to claim that the three-tier application model is "simply outdated". The applications may become large and complex, making it difficult to make frequent changes and adding complexity and ineffectiveness to the maintenance process. The article recommends using the three-tier architecture primarily for small applications, where it may still be the most suitable solution. For larger systems with cloud support, other models such as cloud-based or event-driven serverless architecture should be considered.

3.3 Data, Information, Knowledge and Wisdom

Information systems store, retrieve and process data about real world entities such as places, products, people etc. within the organization [23]. Their role is to organize and process data in order to provide meaningful information to end users and to extend their knowledge and wisdom in the related domain. The precise meaning of terms may vary slightly from publication to publication. Following is a widely accepted adaptation from [31]:

- Data represent properties of an object, events, or their environment. They are products of observation and have no intrinsic meaning until they are transformed into a relevant form.
- Information is derived an inferred from processed data and has the ability to answer simple questions about real-world entities.
- Knowledge is what makes possible to transform information into instructions. It explains how information can be used for the benefit of end users or the company. Knowledge is often referred to by the term *know-how*.
- Wisdom is the ability to increase effectiveness by answering the question *why*. It involves a deeper understanding of the underlying principles.

Other sources, such as [12], merge the definitions of knowledge and wisdom into a single concept. Additionally, terms are often organized in a pyramid structure to represent their relationships, as shown in 3.4.

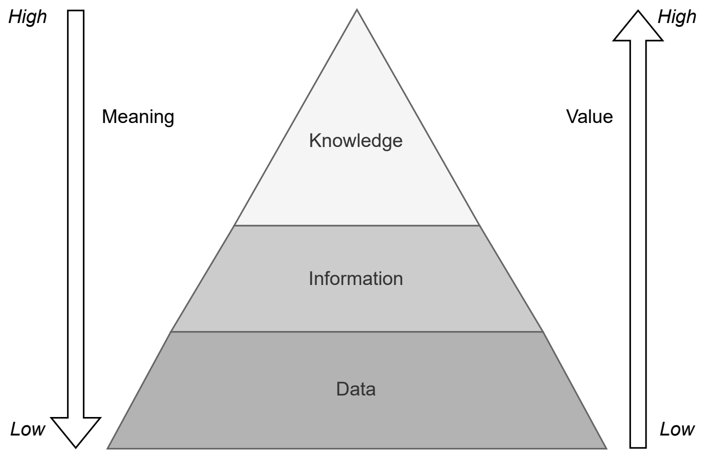


Figure 3.4: The *DIKW* pyramid according to [12]. The value and significance for the company increase with the level of abstraction.

3.4 Exploratory Data Analysis

In general, exploratory data analysis is a data-driven approach which aims to identify patterns and trends in large datasets. Unlike traditional hypothesis-driven methods, which use data as a mean to support or reject a pre-existing hypothesis, the data-driven approach is used to generate hypotheses for future analysis [35].

According to [25], the goal of exploratory data analysis is to provide organizations with actionable insights into the domain of interest. This is achieved by transforming raw data into information that can be used as the basis for making important decisions, which can improve the performance of the organization or company. [25] describes this process in four stages:

1. **Problem definition:** The initial step is to define the problems to be solved and the means to be used to achieve the business objectives. Technical implementation details are secondary at this stage. General questions like „How to reduce hospitalization time“ should be broken down into more specific questions such as „What are the most effective drugs to treat depression“ or „What is the relationship between the type of rehabilitation and the length of stay“. Each question should be answerable using the available data.

2. **Data preparation:** This phase involves the collection and preprocessing of data for further analysis. This typically includes gathering data from various sources, aggregating it, and removing ambiguities, redundancies, and errors. It is a time-consuming step that greatly affects the confidence with which decisions can be made later on.

3. **Implementation of the analysis:** The analysis itself should be performed based on the information obtained in the previous steps. The performed tasks can be divided into three categories, which are described in the following list. Figure 3.5 demonstrates the applicability of different methods depending on the category.

   - **Summarization:** *„Summarization is a process in which the data is reduced for interpretation without sacrificing any important information [25].“* It allows general statements about the data and the population to be made. Summarization is commonly performed using descriptive statistical methods such as mean, median, variance, quartiles, skewness etc.
   - **Finding hidden relationships:** This category involves the identification of data characteristics that may not be immediately apparent from a simple summary and therefore require looking at the data from multiple angles. These characteristics include important facts, relationships, anomalies or trends in the data.
   - **Creating predictive models:** Predictive models are utilized to calculate an estimate for one or more variables based on other variables. They attempt to understand hiddent relationships between its input and output variables. This can be helpful in estimating future events and decision-making. Predictive models are typically divided into two categories: classification and regression models.

4. **Deployment of results:** The final step is to make use of the results of the analysis and apply the solution to the business or scientific problem. The results can be presented as a one-time report or they can be integrated into existing operational processes. For the report, focus should be put on delivering actionable information to people in charge of decisions. In the latter case, the prediction model is required to be able to handle previously unseen data from the system it will be integrated into.

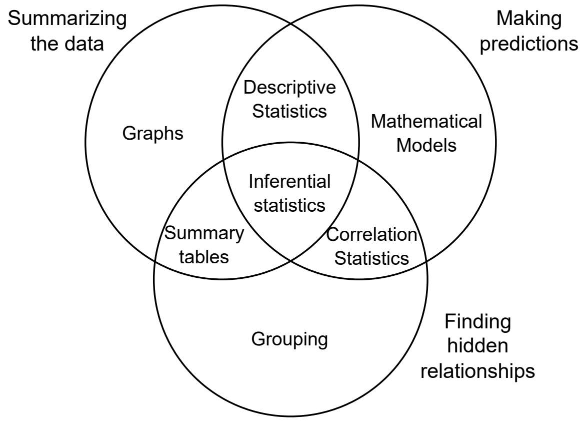


Figure 3.5: The diagram illustrates methods that can be used for data analysis. There is an overlap between the various tasks and the methods that can be used to accomplish them. The position of each method is related to its common usage in addressing specific tasks [25].

Exploratory analysis is an important predecessor of data visualization. It is crucial to develop an understanding of the data and its context in order to gain the ability to effectively communicate the results to the end user.

3.5 Data Visualization Principles

The human brain is capable of processing visual information with great efficiency and the process of data visualization tries to take advantage of this property. Well-designed plots almost always communicate the results more effectively than huge tables or series of numbers and percentages within a body of text. An appropriate figure can help to understand results more accurately [17]. Such visualizations can present data in a way that helps analysts to understand the data and the phenomenon that the data reflect, and to promote analytical reasoning and knowledge building [4]. Psychological studies show that human vision involves abstraction [6]. Seeing actually means subconsciously constructing patterns and extracting high-level features, and it is these patterns and features that analysts use as material for their reasoning [4].

3.5.1 Motivation: The Anscombe's Quartet

One of the most well-known examples demonstrating the importance of graphical representation of data in statistics is *The Anscombe's Quartet*. It was created and published in 1973 by the English statistician Francis John Anscombe in his paper [5] and has been cited in several important publications, including [4, 33, 35, 36]. He created four fictitious datasets consisting of pairs of numeric values. The numerical representation of these datasets is shown in Table 3.1. The summary statistics characterizing these datasets, such as the mean, variance, correlation, and the estimated regression equation, are identical and are shown in Table 3.2. This suggests that the distributions are comparable. However, when plotted into a scatter plot, as depicted in Figure 3.6, it becomes clear that the distributions are completely different. This example illustrates that high-level statistical summaries can be misleading, and that visualization can make the interpretation more precise [4].

Obs. no.	X1	Y1	X2	Y2	X3	Y3	X4	Y4
I	10	8.04	10	9.14	10	7.46	8	6.58
II	8	6.95	8	8.14	8	6.77	8	5.76
III	13	7.58	13	8.74	13	12.74	8	7.71
IV	9	8.81	9	8.77	9	7.11	8	8.84
V	11	8.33	11	9.26	11	7.81	8	8.47
VI	14	9.96	14	8.10	14	8.84	8	7.04
VII	6	7.24	6	6.13	6	6.08	8	5.25
VIII	4	4.26	4	3.10	4	5.39	19	12.50
IX	12	10.84	12	9.13	12	8.15	8	5.56
X	7	4.82	7	7.26	7	6.42	8	7.91
XI	5	5.68	5	4.74	5	5.73	8	6.89

Table 3.1: The Anscombe's Quartet datasets. Each dataset consists of 11 (x, y) pairs. The x-values of the first three datasets are identical [5].

Statistic	Value
Mean of X	9.0
Mean of Y	7.5
Variance of X	11.0
Variance of Y	4.1
Correlation	0.816
Regression Line	$Y = 3.0 + 0.5X$

Table 3.2: A summary of the statistical characteristics of the datasets in Anscombe's Quartet [33]. Given their identical nature, it can be incorrectly assumed that the datasets are also identical.

3.5.2 Gestalt Principles of Visual Perception

Gestalt theory is a psychological discipline developed in the early 20th century by German psychologists that deals with visual perception. The name "Gestalt" is derived from the German word *gestalt* which means *shape* or *form*. One of the key ideas in Gestalt psychology is that the appearance of any element depends on its place and function within an overall pattern [7]. The ideas of Gestalt theory continue to influence numerous disciplines to this day and are particularly useful for creating graphs and visuals that leverage the analyst's visual processing network. Following are the six principles proposed by Gestalt theory [33]:

1. **Proximity:** The brain processes objects that are in close proximity as belonging to a group. Different graphical elements can be grouped together, such as labels with points, bars with each other, or points into clusters [33].

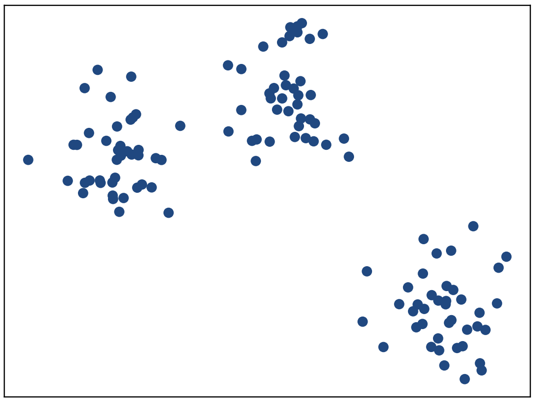


Figure 3.7: Demonstration of the brain's capability to identify the relationship between points within a cluster.

2. **Similarity:** The brain groups together objects that share the same color, shape, size, or orientation [22].

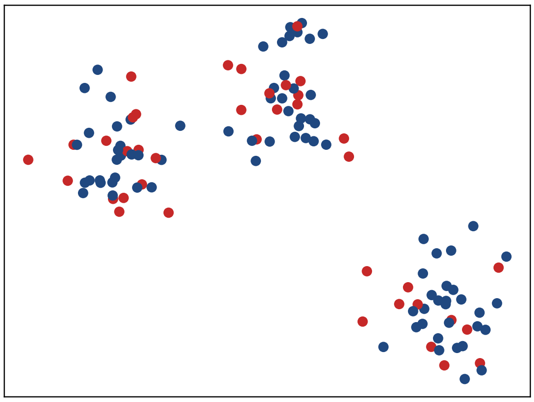


Figure 3.8: Demonstration of the brain's capability to identify the relationship between points across different clusters based on a shared property, in this case color.

3. **Enclosure:** Objects that are physically enclosed together are perceived as belonging into a group. The enclosure can be realized by drawing a border around the group or by coloring the background of the area where the points are located [22].

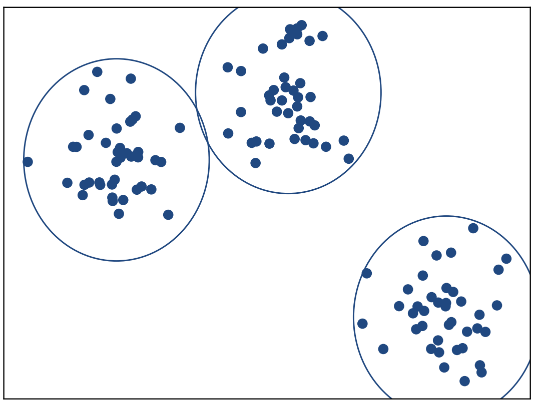


Figure 3.9: Using boundaries to group objects together.

4. **Closure:** The brain exhibits a natural tendency to fill missing information with estimates based on already seen constructs. For example, simply plotting two intersecting perpendicular lines is sufficient for an analyst to perceive the enclosed space of a coordinate plane, even though the complete rectangular borders are not explicitly drawn [33].

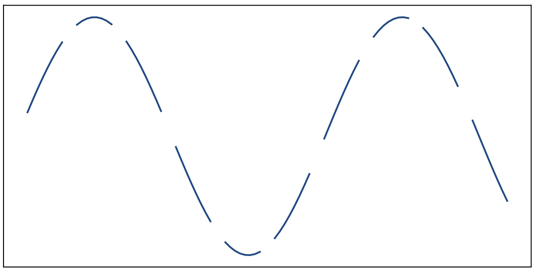


Figure 3.10: Demonstration of the brain's capability to fill in the missing pieces of the plot of a sine function.

5. **Continuity:** Similarly to the principle of closure, the human eye tends to follow the smoothest path when following a sequence of shapes, therefore creating a sense of continuity. This allows to omit certain elements in plots, such as leading lines or axes [33].

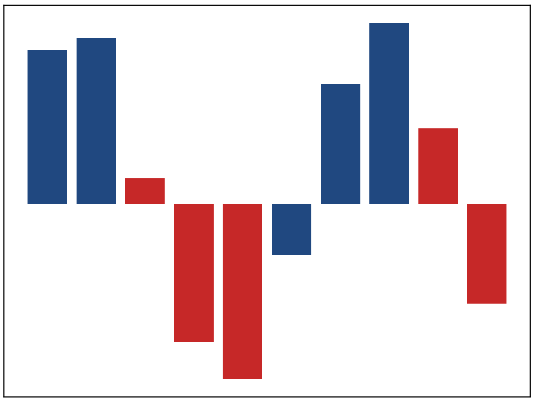


Figure 3.11: The brain knows that the bars are aligned along the x-axis, demonstrating that an explicit x-axis is not necessary for the comprehension of the chart.

6. **Connection:** Objects that are physically connected together are perceived as members of the same group. The connective property typically has a stronger associative value than similar color, size, or shape [22].

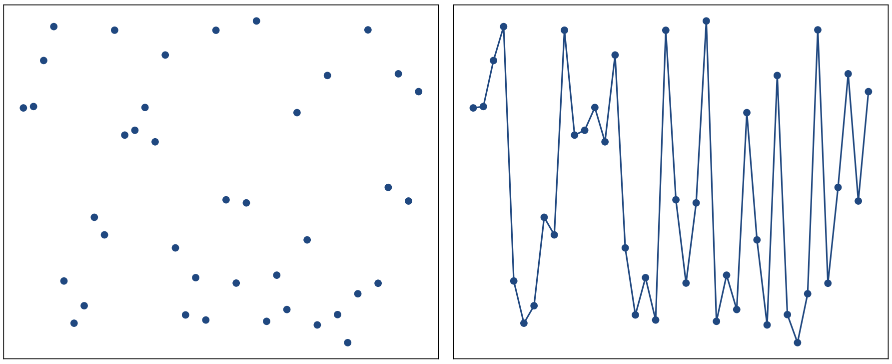


Figure 3.12: The connection of data points helps the brain to see order within the data. By linking individual elements, the brain is able to process complex information more efficiently, discerning relationships and trends that might otherwise remain obscured.

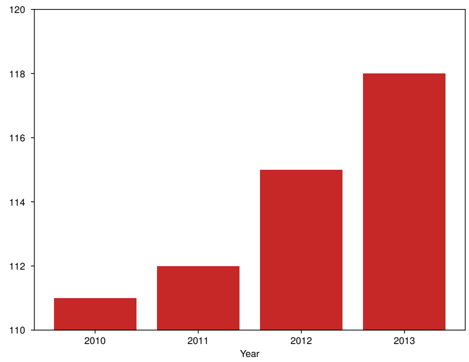
3.5.3 Tufte's Principles of Graphical Integrity

Edward Rolf Tufte is regarded as one of the most influential authorities in the field of data and statistical visualization [43]. In his book [36], he presents a set of principles of visual display that are designed to enhance the clarity and effectiveness of graphical displays of information. These principles are based on the idea that excellence in visual design is largely achieved through the creation of graphics that correspond with the mental tasks they are meant to support [43]. Following is an adaptation of Tufte's principles, as presented by Steven Skiena in his book [35]:

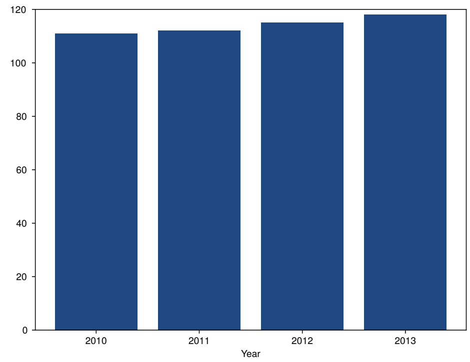
1. **Maximize data-ink ratio:** The primary objective of the visualizations is to show the data. This can be achieved by increasing the *data-ink ratio*, which is defined by the Equation 3.1. Maximizing the data-ink ratio lets data talk by eliminating redundant information [35]. [36] demonstrates the difference by contrasting two line charts created by William Playfair, the pioneer of numerous statistical visualization techniques. The comparison is presented if Figure 3.13.

$$\text{Data-Ink Ratio} = \frac{\text{Data-Ink}}{\text{Total ink used in graphic}} \tag{3.1}$$

2. **Minimize the lie factor:** Visualizations are meant to tell a true story about the data. But improperly designed figures can be misleading and prone to misinterpretations. Potentially harmful practices include presenting only means or interpolations without the actual data, distorting or completely omitting scales, or hiding the origin point [35]. An example of improperly scaled charts is shown in Figure 3.14.



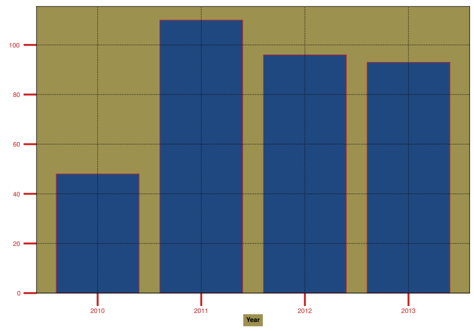
(a) The y-axis of the bar chart incorrectly begins at the value of 110, suggesting a dramatic grow of the y-value over the years 2010 to 2013.



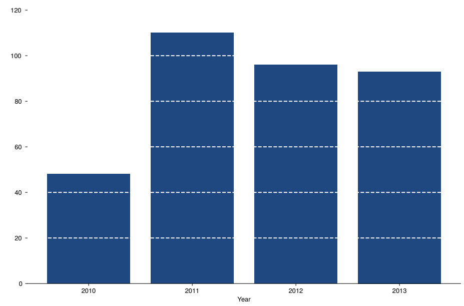
(b) The y-axis of the bar chart correctly begins at the value of 0, reflecting the negligible grow of the y-value over the years 2010 to 2013.

Figure 3.14: An illustration of how modifying the y-axis can impact the interpretation of the data presented in a bar chart.

3. **Minimize chartjunk:** Unnecessary visual elements distract from the core message that a visualization is trying to convey. Charts can often be improved by removing elements instead of adding new ones. Some of the elements that can be omitted without compromising the effectiveness of the visualization are grids, background colors, shadows, or the chart border. It is even possible to utilize the missing ink as a substitute for guiding lines or grids, as demonstrated in Figure 3.15 [35].



(a) The cluttered version of the bar chart with a lot of unnecessary graphical elements.



(b) The clean version of the bar chart after removing distracting visual elements.

Figure 3.15: The difference in readability between a cluttered and a clean bar chart representing the same data.

4. **Use proper scales and clear labeling:** Deficiencies in scaling and labeling of axes represent the primary source of intentional or accidental misinformation in graphs. In order to facilitate the correct interpretation of the chart, accurate labels with appropriate increments must report the proper magnitude and units of numbers [35]. An example of improper scaling was previously presented as part of the *minimize the lie factor* principle in Figure 3.14.

5. **Make effective use of color:** Colors are an important part of any graphical visualization. They play a significant role in charts, serving two primary functions. First, colors can differentiate between classes by assigning distinct colors to elements belonging to different classes. The colors of the classes should be selected to have mnemonic values to naturally link the assigned class. For instance, a chart comparing sources of energy should use green for nuclear power, yellow for solar and black for fossil fuels. Figure 3.8 illustrates the use of colors to distinguish between classes. Secondly, colors can represent a numerical scale. In this case, numerical values are associated with different colors. The scale must be linear, as the reader must be able to distinguish which color comes after the other. A scale can be either fully linear or may distinguish between positive and negative values [35]. Different types of scales are shown in Figure 3.16.

(a) A perceptually uniform sequential scale. Ideal for colorizing data, whose values are always positive. Colormap name: `viridis`.

(b) A diverging scale. Ideal for colorizing data, whose values are both positive and negative. Colormap name: `bwr`

(c) A rainbow scale. This scale is not ideals because it is difficult to comprehend the relative position of each color within the spectrum. Colormap name: `rainbow`

Figure 3.16: An example of color scales available in Matplotlib.

6. **Exploit the power of repetition:** The use of small multiple plots to interpret multivariate data represents an elegant and efficient method of dealing with such data. The cluttering of a single plot with several variables is an inefficient and difficult-to-understand approach. Small multiple charts can be used to break down a distribution and facilitate comparison [35]. An example of such use is depicted in Figure 3.17.

3.6 Chart Types

There are various types of charts that can be used to visualize data. The selection of an appropriate visualization depends on the context. It is the responsibility of analysts to select the most suitable plot for a given dataset and to understand why certain visualizations are more suitable than others [35]. Charts can be used to visualize amounts, distributions, proportions, x-y relationships, geospatial data or uncertainty [42]. Following is a list of different popular chart types and their intended uses [35].

3.6.1 Tables

The simplest and most straightforward way to visualize data is to put them in a table. Although tables may appear to lack visual appeal, they have their place in data visualization and offer various advantages over graphical representations [35].

Advantages

- **Precision:** The values are presented in a precise manner and without rounding errors.
- **Scale:** Given that the numbers in a table are right-justified, they effectively communicate differences in the order-of-magnitude. This property can can be seen in Figure 3.3.

- **Multivariate visualization:** Tables are often the only viable way to visualize data that has a theoretically infinite number of dimensions.
- **Heterogeneous data:** Tables are ideal for presenting a mix of numerical and categorical attributes.
- **Compactness:** Tables are usually the simplest way to visualize a small number of points.

left	center	right
1	1	1
10	10	10
100	100	100
1000	1000	1000

Table 3.3: Right aligning numbers makes comparison easier [35].

There are several ways to improve the design of a table. For example, ordering rows and columns, right-justifying numbers, highlighting cells, avoiding excessively long column labels or adding commas to numbers [35].

3.6.2 Line Charts

When the relationship between two variables can be expressed as a function $y = f(x)$ for a set of (x, y) points, line charts serve as a useful tool for visualizing this relationship. They are especially used in instances when the x values are strictly increasing or when the x axis represents time as shown in Figure 3.18.

Data points in line charts can be connected with straight or smoothed lines that are only intended to guide the eyes, or interpolated with a curve that predicts the values of $f(x)$ for the full range of possible x. In some cases, it may be preferable to omit the connection lines between data points. This is particularly the case when there are many outliers in the data or when the x-axis is discrete [35].

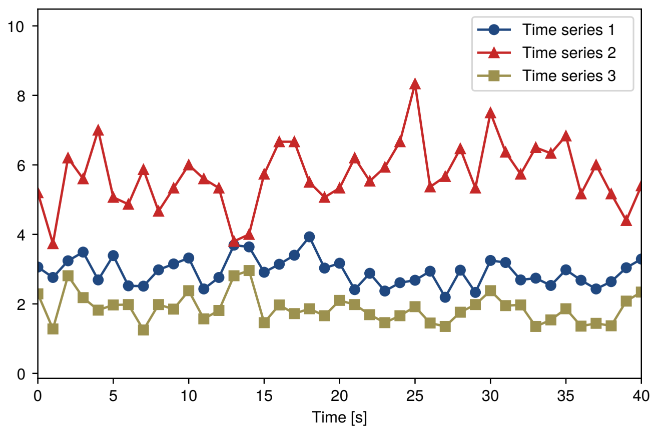


Figure 3.18: An example of a line chart representing values of three series in the span of 40s. The chart makes use of colors and shapes to distinguish between the series. The data were generated randomly.

3.6.3 Scatter Plots

Another chart type used to represent the relationship between variables is scatter plots. These charts display the values of every (x, y) data point in a given dataset [35]. Scatter plots are most effective when used with two-dimensional data. When a third dimension is required, it is often more appropriate to represent it by modifying the color, shape, or size of data points than by adding a third axis. Figure 3.17 illustrates the relationship between different attributes of wine based on their class, using scatter plots.

3.6.4 Bar Plots and Pie Charts

Bar plots and pie charts are tools for presenting the relative proportions of categorical variables [35]. However, visualizing these proportions can be challenging, particularly when the whole is broken into many different pieces or when it is important to clearly observe changes in proportions over time or across conditions. There is no single ideal visualization that always works [42]. Table 3.4 provides a summary of the advantages and disadvantages of different types of visualization.

Property	**Pie chart**	**Stacked**	**Side-by-side**
Clearly visualizes the data as proportions of a whole	✓	✓	✗
Allows easy visual comparison of the relative proportions	✗	✗	✓
Visually emphasizes simple fractions, such as $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$	✓	✗	✗
Looks visually appealing even for very small datasets	✓	✗	✓
Works well when the whole is broken into many pieces	✗	✗	✓
Works well for the visualization of many sets of proportions or time series of proportions	✗	✓	✗

Table 3.4: The advantages and disadvantages of the most common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars [42].

Pie charts are often the subject of criticism from data analysts due to their tendency to occupy more space than necessary and to present data in a less readable and comparable manner [35]. In his book [36], Edward Tufte states, "A table is nearly always preferable to a simple pie chart; the only worse design than a pie chart is several of them." A sentiment echoed by author and analyst Cole Nussbaum Knaflic who, in her book [22], dedicates a chapter to the reasons to avoid pie charts, entitled "Pie charts are evil." However, other sources, such as [35, 42], admit that there are instances when pie charts can be beneficial in data visualization. Figure 3.19 demonstrates one such instance and compares various types of visualizations of proportional data.

3.6.5 Histograms and Density Plots

Histograms and density plots represent an effective approach for visually representing data distributions. They help to understand how a specific variable is distributed within a dataset, answering questions like: Where is the peak of the distribution? Is the distribution symmetric or skewed? How many local modes exist [35]? However, both require parameter choices, which, when made incorrectly, can result in misleading plots [42].

Histograms are generated by binning the data and their exact visual appearance depends on the choice of the bin width. If the bin width is too large, smaller features in the distribution of the data may disappear. Alternatively, if the bin width is too small, the histogram becomes peaky and the main trends in the data may be obscured [42]. An example of the impact of bin width size is depicted in Figure 3.20.

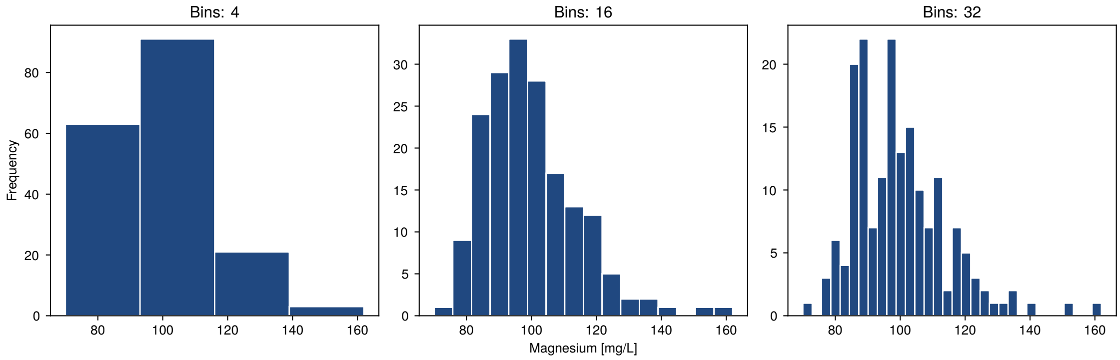


Figure 3.20: An illustration of the effect of the bin width on the properties of the dataset that are revealed. Dataset: [1].

Density plots, on the other hand, attempt to visualize the data's underlying probability distribution function by rendering a continuous curve as shown in Figure 3.21. The estimation of the curve is typically performed using a method known as *Kernel Density Estimation*, which relies on two main parameters: the kernel and the bandwidth. The bandwidth parameter has the same properties and impact as the bin width in histograms. The choice of the kernel affects the shape of the density curve. For instance, a Gaussian kernel will have a tendency to produce density estimates that resemble a Gaussian distribution, with smooth features and tails. Regardless on the parameters, the area under the curve should always equal to one. One potential issue with density plots is their tendency to suggest the presence of data where none exists, particularly in the tails of the distribution [42].

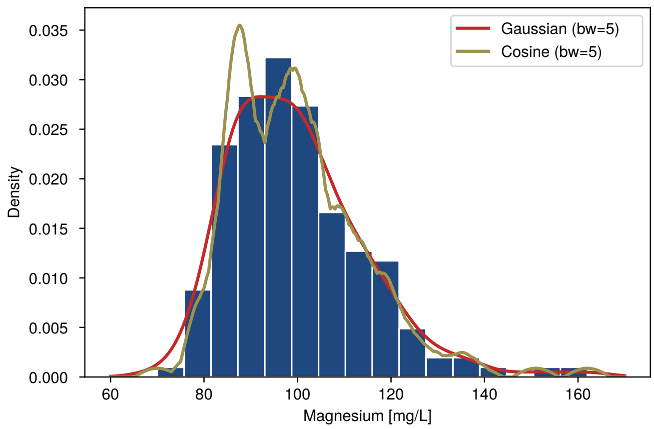


Figure 3.21: An illustration of the estimation of the underlying probability density function of the data. The estimations are based on a Gaussian and a Cosine kernel with a bandwidth of 5. The graph shows that both estimated functions predict the presence of data in the tails, where no real data was recorded. Dataset: [1].

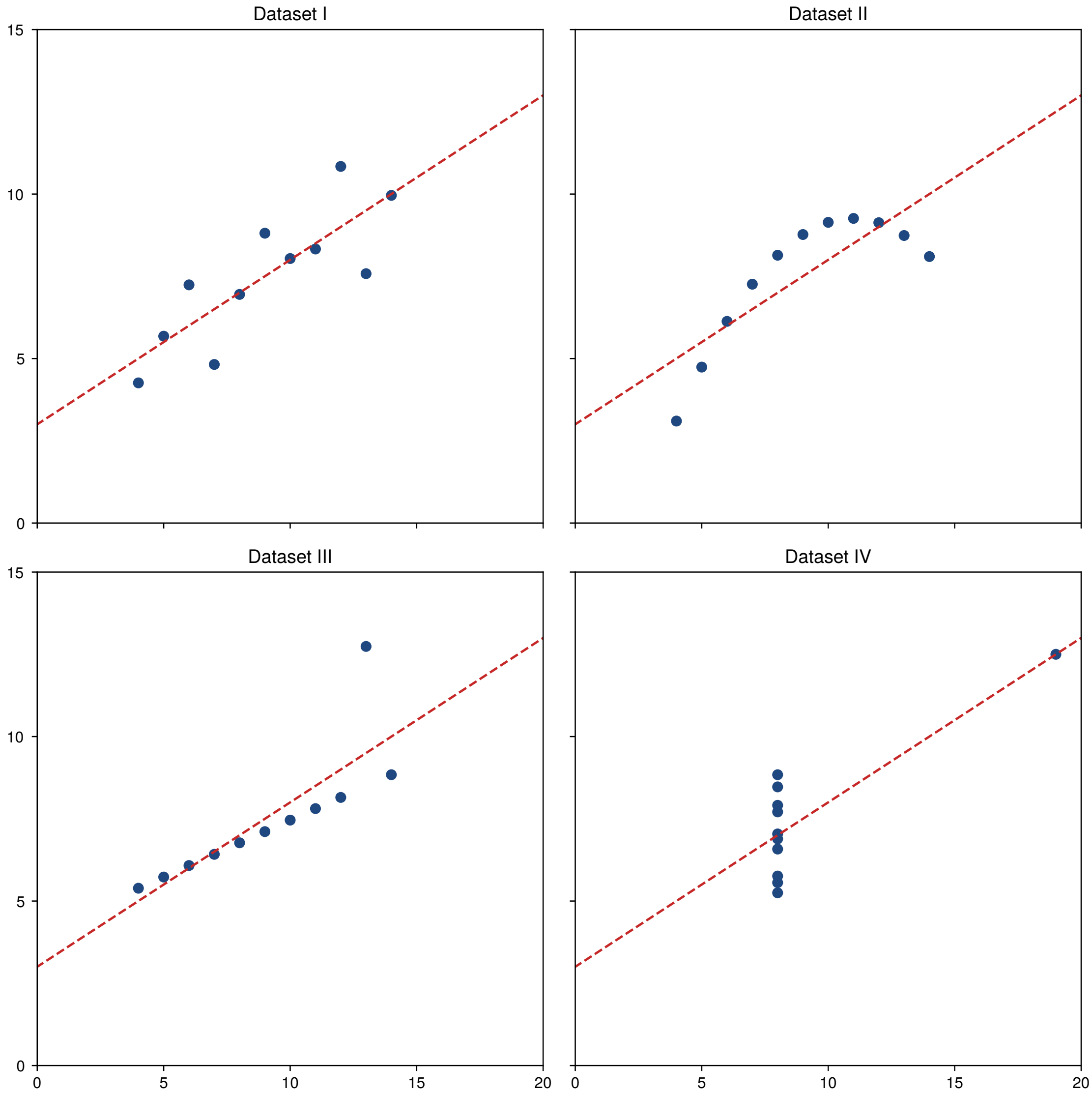
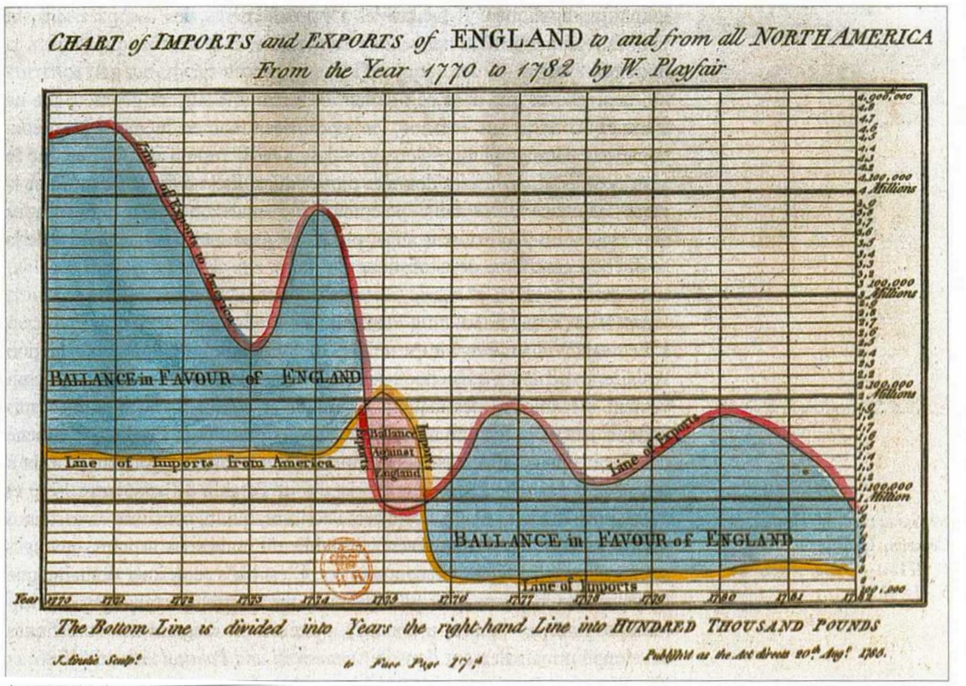
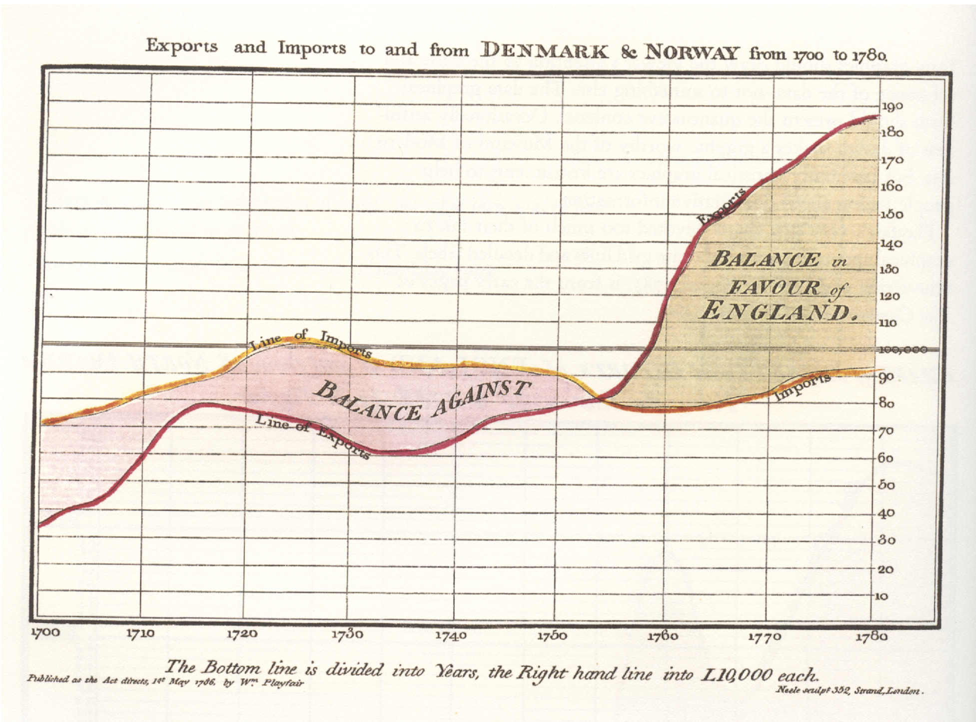


Figure 3.6: A visual representation of the datasets in Anscombe's Quartet [5]. The distinct characteristics of the datasets are clearly evident when presented in a graphical format. Without such a representation, the differences between them would remain concealed.



(a) Too much ink is devoted to grid lines and labels, distracting the reader from the essence of the data [36]. Image source: [26].



(b) The design is cleaner and focuses the attention on the time-series itself [36]. Image source: [26].

Figure 3.13: The evolution of Playfair's line charts, which illustrated the balance between imports and exports between England and various countries. The improvement in graphical design underscores the importance of showing the data over all other design elements [36]

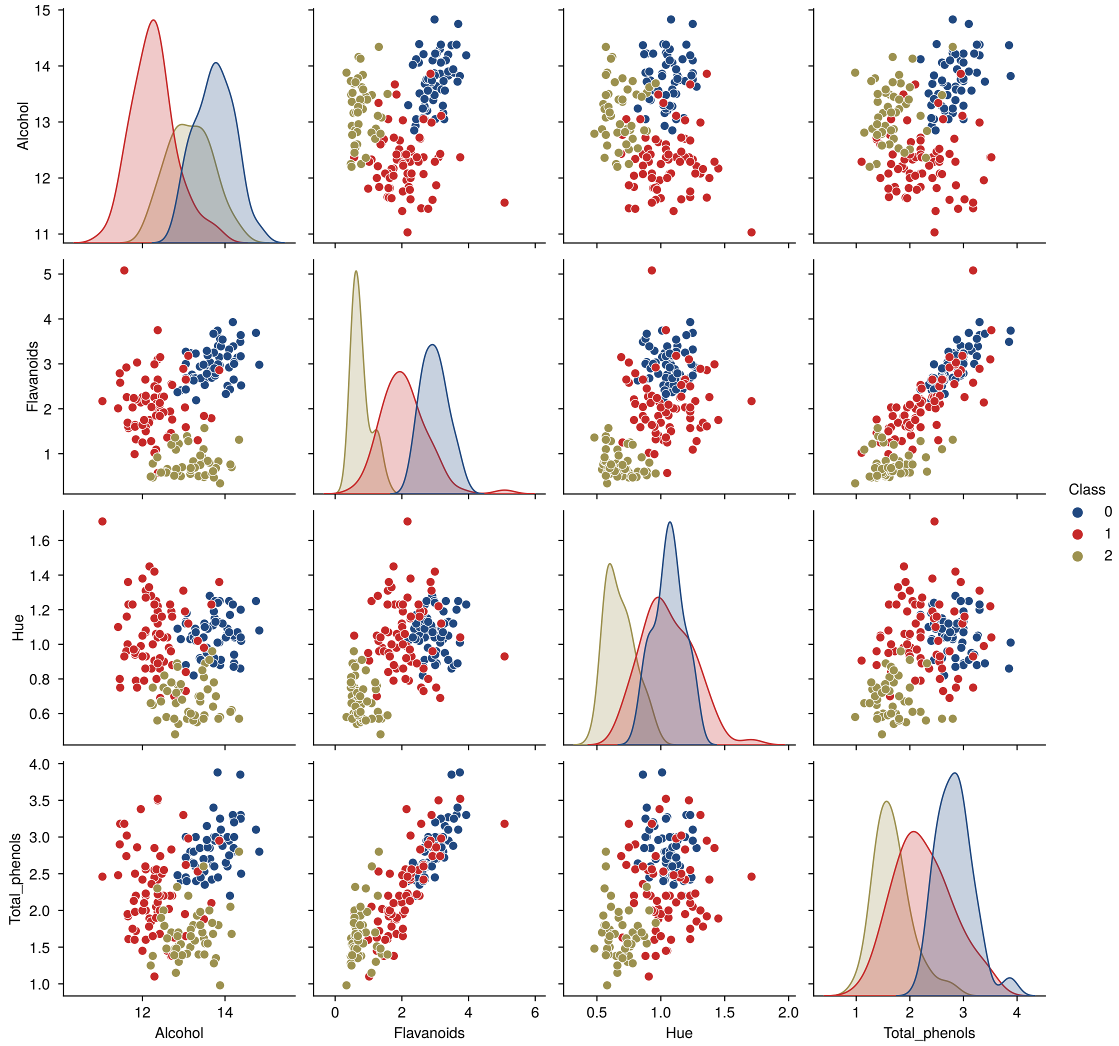
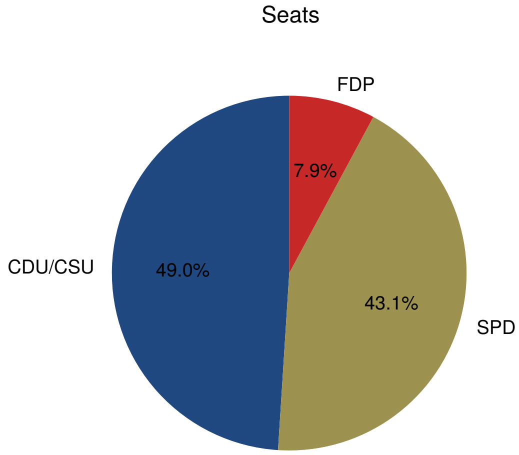
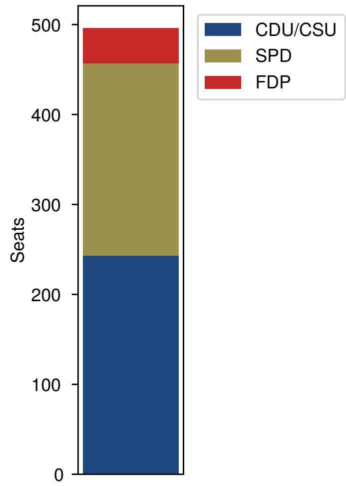


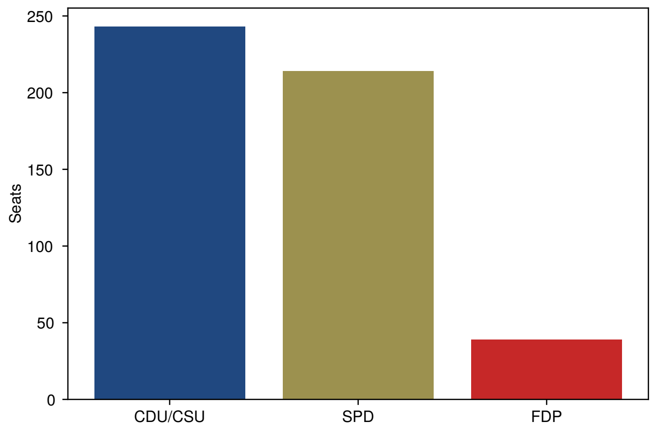
Figure 3.17: An example of the usage of small multiple plots to reveal the distributions of various attributes of wines and the relationships between them. Dataset: [1].



(a) The pie chart is the only visualization that clearly shows that the SPD and FDP together hold a majority.



(b) A stacked bar makes it easier to see the total number of seats.



(c) Plotting the data using side-by-side bars is the best approach to allow for direct comparison between groups.

Figure 3.19: A comparison of different approaches to visualizing proportions. The charts represent the party composition of the German parliament between the years 1976 and 1980. The CDU/CSU held 243 seats, SPD 214, and FDP 39, for a total of 496 seats [42].

Chapter 4

Analysis of Requirements

This chapter defines the goals and requirements for a new information system intended for internal use at the Rehabilitation Center Kladruby, hereafter referred to as "RCK". The requirements emerged from multiple consultations with representatives from the center, specifically clinical psychologists PhDr. Petra Fiľová, Ph.D., and Mgr. Pavel Viktora. The RCK, described in further detail in Section 4.1, is a leading rehabilitation facility in the Czech Republic.

The daily volume of administered psychological tests at the center is substantial and requires a significant amount of work to process. The analysis of the results of examinations is currently time-consuming and complicated. The center relies on outdated methods and the way in which data are stored and processed is unfavorable. These practices contribute to an inefficient workflow and diminish the potential quality of data analysis. The details of the current workflow and the proposed new workflow are described in Section 4.2.

The population of patients at the RCK represents a statistically significant sample in the context of both the Czech Republic and Europe. Currently, the data are not efficiently utilized. The implementation of an information system will enable the creation of a digital database of results, which will potentially lead to different scientific research projects, as discussed in Section 4.3.

Section 4.4 describes the format in which tests and norms should be represented within the system. Following is Section 4.5, which addresses the examination results and the manner in which the system should store and present them. Both visualizations and exports are discussed in this section.

The last section, Section 4.6, describes the system's requirements that emerged after a proof of concept prototype was tested with representatives of the RCK. The section also touches on the user interface of the system.

4.1 Rehabilitation Center Kladruby

The Rehabilitation Center Kladruby is a specialized medical institute and one of the leading state centers in the Czech Republic. It is located in the small village of Kladruby, approximately 60 km southeast of Prague. The center was established in the late 1930s, originally serving as a sanatorium. During the war the institute served as a German military hospital and after the war it became a Soviet repatriation station. In 1946, the Ministry of Health of the Czech Republic took over the facility [29].

The center provides complex inpatient rehabilitation care for patients with a wide range of disabilities of varying degrees, especially neurological and musculoskeletal disorders, post-surgical conditions, spinal injuries, and other medical conditions. For the most sever cases it also focuses on their resocialization and smooth reintegration into society. The center also participates in the education of students of medical and social schools, medical, rehabilitation and health and social faculties and in the education of other professionals working in the health care sector [30].

As of 2022, the center had a capacity of 275 beds, with a new building under construction to increase the capacity to approximately 300 beds in 2025. The medical team comprised 236 specialists. The yearly average number of patients who underwent treatment at the center between 2015 and 2022 was approximately 1700. In 2022, the average length of stay was 56 days [28].

4.2 Workflow

The process of administering and evaluating psychological or neurological tests is complex and lengthy. Clinical psychologists must first understand the patient's issues and medical condition to select the most suitable test, as the choice has a significant impact on the informative and explanatory value of the examination. Conducting these tests requires considerable time, concentration, and expertise from clinicians. The results need to be measured precisely and stored in an organized manner. However, the current data management and storage practices at the RCK are suboptimal. General information about patients, the results of their examinations, and the norms used to evaluate the results are spread across different data storages, which further complicates the clinicians' workflow.

Once clinicians complete the examination, they need to thoroughly evaluate the results. Psychologists primarily refer to norms to put the score into the context of the reference group, which is determined by various factors, such as age, sex, diagnosis etc. However, norms are not always precise, and they may be created for different reference groups. Moreover, the center relies on outdated methods, such as manual inspection or spreadsheets, which are inherently ineffective. In addition to the time inefficiency of this approach, the manual or semi-automated calculation of normed scores (discussed in Section 2.4) may introduce an error in the resulting scores due to the potential for human error.

Clinicians may also need to use alternative methods of comparison. Given the considerable number of patients at the RCK, it is beneficial to compare the results with historical data from similar examinations on other patients. Another crucial indicator of a patient's performance is their historical evolution of the results. This allows clinicians to clearly identify whether the patient's cognitive abilities are improving or declining. In certain instances, psychologists may also benefit from grouping multiple examinations of a single patient together, based on the cognitive functions that they test. Currently, clinicians lack the flexibility to evaluate patients efficiently as some comparisons are difficult or impossible to perform.

The digitization process is expected to result in several improvements over the current state. The primary objective is to reduce the time required for the testing process. A centralized digital storage of exams will simplify the process of sharing results among psychologists and eliminate the error-prone step of manually calculating normed scores. The new workflow, shown in Figure 4.1, will allow clinicians to focus more on the results and easily create graphs and summary views.

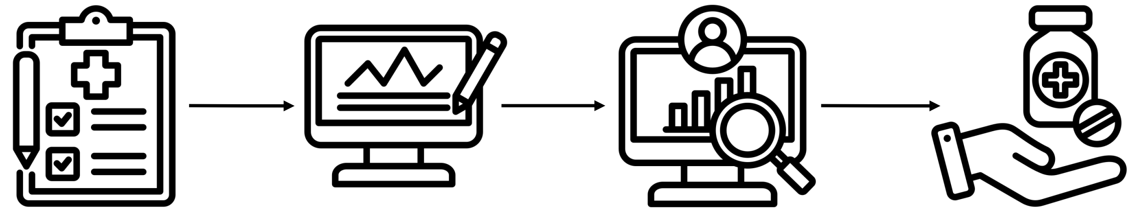


Figure 4.1: The digitalization of the process of administering psychometric tests is expected to increase its efficacy and to provide psychologists with additional support in their decision-making.[1]

4.3 Research Goals

In order to draw useful inferences about patients' relative performance, clinicians use norms to place the result into perspective with different comparison groups. The validity of the interpretation of this results is dependent on the appropriateness of the reference group. It is important that reference populations be carefully defined and clearly described. The interpretation of such results is also dependent on the accuracy with which norms summarize the performance of the reference population. Furthermore, the usefulness of norms based on a given sample may diminish over time. Consequently, for tests that have been in use for a substantial number of years, it may be necessary to renorm to ensure the continued validity of norm-referenced test score interpretations. It is possible for more than one reference population to be appropriate for the same test. For instance, test performance might initially be interpreted in relation to local norms based on sampling from a particular facility, before being compared with the national reference group using national norms. In other cases, norms might be based on diagnostic or educational classifications [2].

It can be argued that the current state of psychodiagnostics in the Czech Republic is suboptimal [37]. Some norms lack precision or clinical utility, as in the case of the Cognitive Estimations Test, whose norms are not considered by the center. The accumulation of data in digital form will serve as a basis for creating new, standardized neuropsychological norms for patients who have suffered a spinal cord or brain injury. Furthermore, a digital database will provide data that can be used for further research, potentially benefiting students of psychology and medicine fields at Masaryk University in their bachelor's and master's theses (FIĽOVÁ, Petra. Apr 18, 2024. Private communication). A diagram of the steps involved in the process and the possible outcomes is shown in Figure 4.2. The establishment of a digital database from examinations performed within Rehabilitation center Kladruby represents a significant promise of gaining access to a statistically robust sample of data on the specifics of the level of cognitive functions of patients. This sample is exceptional not only in the context of the Czech Republic but also in Europe and will contribute to numerous research projects in the field of neuropsychology (FIĽOVÁ, Petra. Apr 18, 2024. Private communication).

[1] Icons source: https://www.flaticon.com/

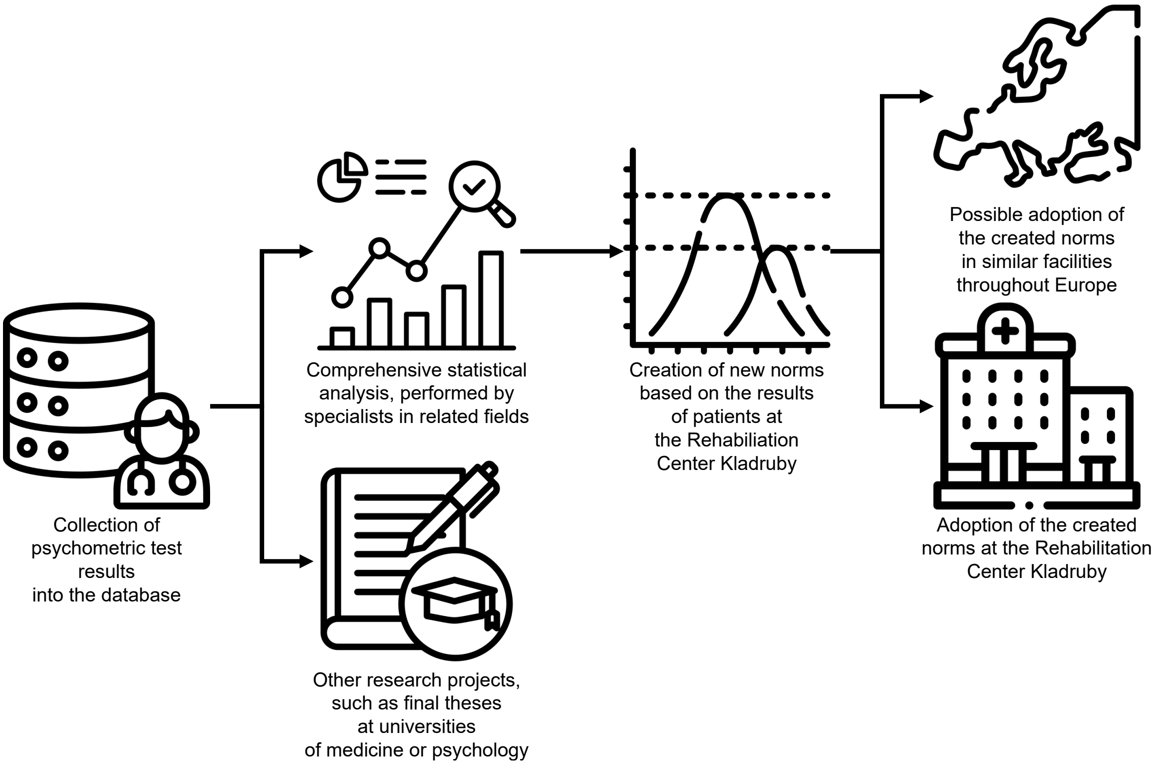


Figure 4.2: A digital database of results is expected to serve as fertile ground for a range of various research projects.[2]

4.4 Test Management

One of the fundamental capabilities of an information system serving as a database of examination results is the ability to store and manage information about tests and their norms. The RCK primarily utilizes intelligence and neuropsychological tests. Appendix C provides a list of some of the tests that are currently intended to be digitalized. It should be noted that the list may be subject to future expansion.

Beyond basic information such as the test's name, abbreviation, and description, a test is represented by a table with flexible dimensions $m \times n$, each having specific row and column names. A test can also include a theoretically unlimited number of subtests, which are similarly structured as tables with variable dimensions. Both tests and subtests include information about the cognitive functions they are designed to assess. As clinicians will only upload the results into the database upon completion of the examination, it is unnecessary to store the entire assessment protocol in the database. Clinicians will continue to primarily utilize physical guidelines as reference material.

The tests intended to be used within the information system are norm-referenced, meaning that they use norms to interpret and compare the test score of a given examinee with those of others [2]. More precisely, they use standard scores, which were introduced in Section 2.4. The possible number of norms associated with each test is unlimited. Norms define reference values for all the cells of a test. They also include information about the

[2]Icons source: https://www.flaticon.com/

age range, premorbid education, and gender of the reference group. Norm values are either defined as parameters of a Gaussian distribution or as look up tables. In the first case, as illustrated in Table 4.1, each cell is assigned a mean and standard deviation. These parameters are employed to calculate the z-score from the raw result using the formula presented in equation 4.1 [3]. In the latter case, multiple intervals of raw scores are defined and associated with normed values. This second type of norm definition is utilized as a lookup table, as depicted in Table 4.2. Both tables 4.1 and 4.2 are structured in a way that is often seen in psychological literature. The methods of defining norms in psychology presented here are not exhaustive, but they cover the current needs of the RCK.

WAIS III [Age: 16–25, Sex: Male]		
Subtest	**Mean**	**Standard Deviation**
Picture Completion	50.5	3.2
Similarities	63.8	4.8
Block Design	14.1	1.1
Digit Span	78.9	3.4
Symbol Search	33.0	1.8
...		

Table 4.1: An example of a norm definition for the Wechsler Adult Intelligence Scale III test. The z-score of a patient for a given subtest is calculated using the associated parameters of the normal distribution of the reference group, in this case for male patients in the age range of 16 to 25. All the values in this table are intended for demonstration purposes only and were generated randomly.

$$Z = \frac{X - \mu}{\sigma} \tag{4.1}$$

where:

- $\boldsymbol{Z}$ is the Z-score of patient.
- $\boldsymbol{X}$ is the raw score of the patient.
- $\boldsymbol{\mu}$ is the mean (average) of raw scores in the reference population.
- $\boldsymbol{\sigma}$ is the standard deviation of raw scores in the reference population.

WAIS III - Picture Completion [Sex: Male]					
Age / Weighted score	16–25	26–35	36–45	46–55	
1	54–56	54–56	53–55	52–54	⋮
2	57–59	57–59	54–56	55–56	
3	60–63	60–62	57–58	57–57	
4	63–66	62–65	59–61	58–61	
...					

Table 4.2: An example of a norm definition for the Picture Completion subtest of the Wechsler Adult Intelligence Scale III test. The weighted score is calculated by looking up the table for the raw value. For example, if a male patient in the age range between 36 and 45 scores 58 points in the subtest, the weighted value of his score is 3 points. All the values in this table are intended for demonstration purposes only and were generated randomly.

The format in which the norms are available at the RCK poses a challenge to the process of uploading the norms into the system. The majority of the norms are stored in physical form or PDF scans that lack a consistent format. Tables have varying dimensions, captions, and structures. This makes the automation of the process difficult and creates an opportunity for future projects in this area.

4.5 Test Result Analysis

Clinicians measure the results of conducted tests as raw numerical values. They then need to convert the raw values to normalized values according to a carefully selected norm. The conversion is often time-consuming and error-prone for two reasons. First, the test guidelines are dense and reading in them requires a substantial amount of attention and precision. Secondly, clinicians are required to execute numerous manual numerical operations, which inherently increases the likelihood of mistakes. It is crucial that the system implement a mechanism that automates the conversion of raw scores into normed scores, and that it allow for both to be easily displayed. The results should be visible both as numerical values in tabular format and graphically as markers on probability distribution charts.

4.5.1 Visualizations

An interpretation based solely on numerical values of normalized scores would be incomplete, and therefore clinicians rely on additional and more precisely tailored comparisons. These comparisons are well-suited to be visualized graphically. A set of visualizations was discussed with the representatives of the RCK, along with their benefits in the context of analysis. The proposed visualizations will directly improve the workflow discussed in Section 4.2.

Comparison with Reference Group

First of all, clinicians need compare the results of a single patient with those of other patients at the clinic. They need to precisely select the comparison group by filtering based on multiple criteria, such as gender, age, education level prior to condition, type of injury, and diagnosis. This process enables clinicians to identify factors that have the most significant impact on patient results.

Historic Development

Another method that helps clinicians to comprehend more deeply the pathology of the patient is the visualization of the historical development of the patient's results. It is important to include the results of the entire test and all associated subtests in a single view. The visualization helps assess the effectiveness of the medical treatment used by providing a direct comparison of the cognitive abilities of a patient in time.

Comparison between Tests

A third visualization was proposed and discussed during meetings with representatives from the RCK. This visualization presents the results of a single patient across all tests that assess the same cognitive function. This was thought to provide a more comprehensive view of the patient's cognitive abilities. However, comparing results between different tests is problematic. Due to the doubtful usefulness and the lack of precise specifications, the discussion of this visualization was postponed. It may be considered for a future upgrade only if its utility is proven.

4.5.2 Exports

As previously discussed in Section 4.3, one of the objectives is to create a database of examination results that can be utilized in research. For this reason, the data must be easily retrievable from the database by predefined exports. The exports should support formats that are easily shareable and widely understood by a broad community of researchers. In line with current requirements, two types of data exports have been considered.

Export of Test Results

In the context of research, the raw results are the most valuable aspect of the collected data. This export will provide researchers with a dataset that will serve as the foundation for establishing new norms. The data contained in this export should include all the examinations associated with a specific type of test. Each record must contain all information necessary for the interpretation of the examination, such as raw results, the patient's age on the day of the examination, and the patient code. The records should include all the examinations of all the patients in the database.

Export of Patient Details

The second export will serve to complement the previous one by providing additional patient-specific information. It should include a list of all patients and their respective parameters. This includes the patient's code, date of birth, gender, education level prior to condition, types of injury, and assigned diagnoses.

4.6 System Requirements

The system is mainly intended to be used by a small group of clinical psychologists and neuropsychiatrists at the RCK. Access will also be granted to mental health professionals or trainees in their pre-certification process, provided they submit a certified copy of their specialized training in health care. This type of access will be restricted and managed by the RCK. Users may vary in age and technical skills. The majority will possess educational backgrounds in non-technical disciplines, such as psychology or medicine. The characteristics of the user base are closely correlated with the requirements for the system's user interface, which is discussed in greater detail in Subsection 4.6.1.

The nature of the target user group complicates the process of precisely specifying the system's requirements due to the knowledge gap between developers and psychologists. The initial requirements were vague and imprecise. For this reason, a proof of concept prototype was developed. This prototype, detailed in Subsection 4.6.2, improved communication with the representatives of the RCK and enabled a more detailed definition of system requirements. The feedback on the prototype led to a list of enhancements that had not been previously considered, as described in Subsection 4.6.3.

4.6.1 User Interface

There were few specific requirements regarding the user interface provided by representatives of the center. It was expressed that the user interface should be intuitive and easy to understand. As the primary platform for users to access the system will be desktop computers and laptops, it was recommended that space be utilized efficiently on 16:9 screens. A column layout that would visualize both the details of a patient and the results of a test was suggested to avoid the need to constantly switch pages.

Considering the typical user profile, the system's interface should be intuitive enough to not require external explanations for users with an education in psychology and related fields. This objective may be achieved through the implementation of a straightforward design complemented by help buttons.

4.6.2 Prototype

A prototype that served as a proof of concept solution was developed during the autumn of 2022 by Bc. Miroslav Šafář, Bc. Matěj Mudra, and Bc. Lukáš Foltyn. The application was part of an assignment within the course *Information Systems* at Brno University of Technology. The prototype's architecture comprised a REST API in .NET 6[3] for the backend and a Flutter-based[4] web application for the frontend. The user interface of the prototype is presented in Appendix B. It included the fundamental features that had been originally requested by the RCK. The prototype implemented user authentication with roles and management modules for users and patients. The system enabled users to create tests and define a single norm per test. Norms could be defined exclusively as parameters of a normal distribution. The results of examinations were only available in raw form. The prototype also allowed patients to be assigned multiple diagnoses. Although the prototype served as a solid foundation for early testing, it was recognized as lacking several supplementary features.

4.6.3 Detailed Requirements

A part of the process of specifying requirements involved testing and discussing the prototype with PhDr. Petra Fiľová, Ph.D., and Mgr. Pavel Viktora. This allowed the representatives to provide more detailed feedback and led to a greater level of detail in the

[3] https://dotnet.microsoft.com/
[4] https://flutter.dev/

specifications. In addition to the general requirements described in the previous sections of this chapter, the following requirements were introduced:

- Tests should support multiple norm definitions.
- Norms should include attributes that specify the reference group, including age, sex, and education.
- Support for normed scores should be implemented, allowing for quick conversion between raw and normed scores.
- The definition of standard score norms using intervals should be possible.
- The system should incorporate support for categorizing diagnosis groups.
- Rehabilitations should include start and end dates.
- Tests should include attributes that detail the cognitive functions they assess.
- Examinations should include an attribute to automatically track the sequence number of examinations.
- A feature should be available to track a patient's diagnosis history.
- Patient attributes should include premorbid education levels.
- Patient attributes should include the date of hospitalization.
- Patient attributes should include a set of keywords that clarify the extent of an illness or brain damage.

Chapter 5

Design

This chapter describes the design of an information system that aims to fulfill the objectives defined in sections 4.2 and 4.3 of the previous chapter. Section 5.1 provides an overview of the adaptation of the three-layer architecture used as the base of the system. A user role model, comprising three distinct roles, is employed to manage system access. The exact actions that each role is allowed to perform are presented in Section 5.2. The data layer of the solution is represented by a complex data model, which is described in detail in Section 5.3. The interface of the REST API, which powers the application layer, is introduced in Section 5.4. The presentation layer is completely decoupled from the underlying layers and is represented by a web application, which is described in Section 5.5. The last part of the chapter, consisting of Sections 5.6 and 5.7, covers the designed visualizations and exports, respectively.

5.1 Architecture

The system uses a variation of the three-tier architecture described in Section 3.2. This architecture is robust enough to meet the requirements for security, connectivity, and scalability of the solution while avoiding the complexity associated with cloud architectures. Figure 5.1 illustrates the structure of the proposed architecture. The application layer and data storage layer compose the backend of the system, while the presentation layer represents the frontend. The following list expands on the properties of each layer:

1. **Presentation layer:** The presentation consists of a web application, as this allows the client to be deployed to any computer with a web browser, thus facilitating the deployment process.

2. **Application layer:** This layer comprises a REST API that exposes a set of endpoints for data manipulation within the database.

3. **Data storage layer:** The data storage layer is represented by a relational database server. During the design process, both SQL and NoSQL databases were considered as potential solutions. Ultimately, the SQL database was selected for the following reasons: the data contain a significant number of relationships that are not well handled by NoSQL databases. Moreover, SQL databases ensure better data consistency, which is an important property in the context of medical information systems. Finally, given the relatively small amount of expected data, horizontal scaling is not necessary.

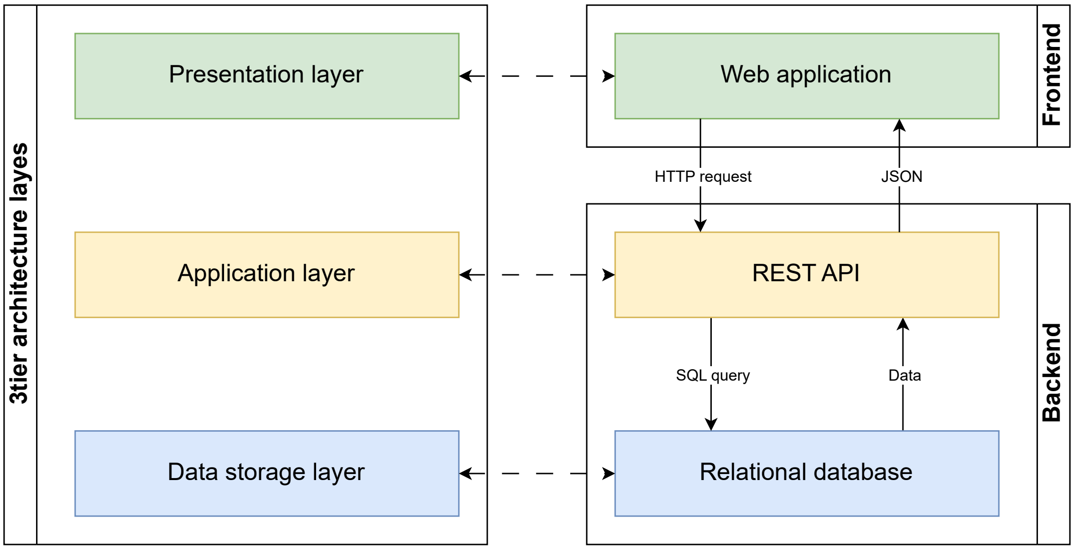


Figure 5.1: The architecture of the proposed system is based on an adaptation of the three-tier architecture. The REST API serves as an interface for clients to manage data in the database.

5.2 User Management

The users who will interact with the system can be classified into three categories, each requiring distinct levels of access rights. Therefore, three user roles have been introduced. Table 5.1 summarizes the permissions associated with each role. The specific actions that the users are permitted to perform are further illustrated in the use case diagram in Figure 5.2, which serves as a complement to the table of permissions. In order to gain access to the system, users must first be provided with an account and credentials in the form of a username and a password. This process is the responsibility of the administrator, who has complete control over the user base that has access to the system.

Role	**Permissions**
Clinician	View and manage patients and assigned examinations, diagnoses, rehabilitations, and damage specifications. View tests, norms, cognitive functions, diagnoses, and damage specifications.
Moderator	Create tests and manage norms, cognitive functions, diagnoses, and damage specifications.
Admin	Manage user accounts.

Table 5.1: Users are allowed to view and manage different database entities based on the role assigned. The roles form a hierarchy.

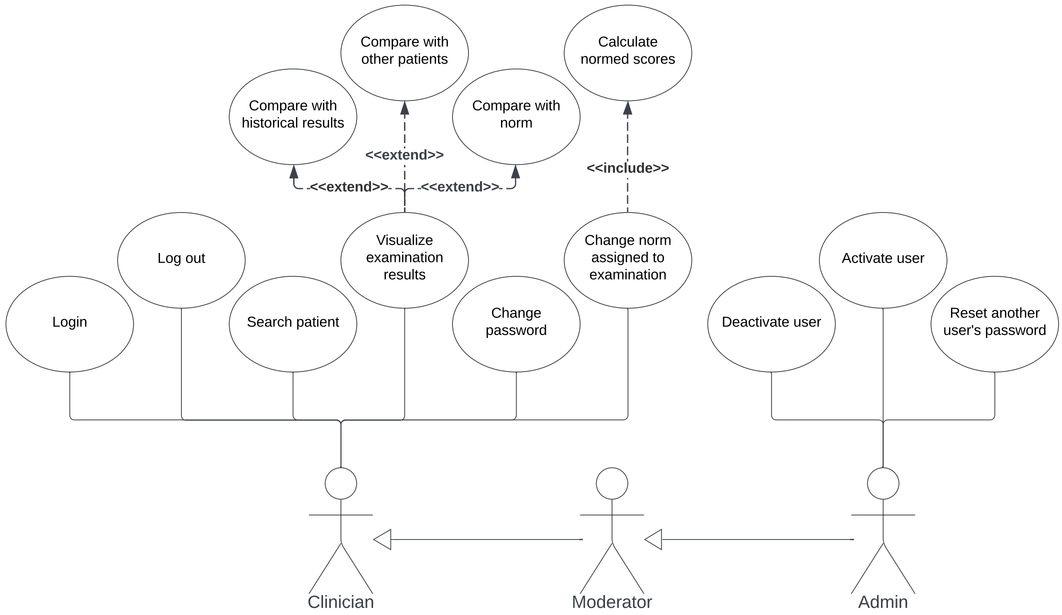


Figure 5.2: The use case diagram of the system. The use cases are not exhaustive, as operations that directly manipulate database entities were omitted for the sake of simplicity. These operations are presented in Table 5.1.

5.3 Data Model

The designed data model consists of 21 entities, six of which represent many-to-many relationships. An entity-relationship diagram, presented in Figure 5.3, visually demonstrates the schema of the model. The model supports general definitions of tests, as each test can have an unlimited number of rows, columns, and subtests. The relationship between a test and a subtest is recursive, allowing for an unlimited hierarchy of nested subtests. These properties exceed the requirements of the Rehabilitation Center Kladruby (RCK) and are designed for future-proofing the model. Following is a list of the real world concepts that the system exposes to users and the data model entities underlying these concepts:

- **User:** The concept is directly mapped to `User` entity. In addition, each user is assigned a role, which defines their permissions. This is represented by the entities `Role` and `Permission`. The `Refresh Token` entity enables users to extend their sessions without re-authenticating.
- **Patient:** This is a representation of a patient in the real world. The concept is simply represented by the `Patient` entity.
- **Test:** Tests are represented by multiple entities. The first entity, `Test`, contains general information about the test, such as its name and abbreviation. Entities `Test Row` and `Test Column` store details about each row and column, including their name and position in the table. Finally, the relationship with `Cognitive Function` assigns a list of assessed cognitive functions to the test.

- **Examination:** An examination comprises both raw and normed scores for all the cells of the test assessed. The entities underlying this concept are `Medical Examination`, which represent general information about the examination, and `CellResult`, which is used to store the results of both raw and normed results for each cell of the test. It is necessary to store the normed scores in the database because they are needed for the generation of higher-level visualizations.

- **Norm Definition:** Norms are represented by two entities. `Norm Definition`, which stores general information about the norm, and the relationship `CellNorm`. In contrast to examination, where only one value is stored for each cell, norm definitions may require to store more than one value. For this reason, the entity `CellNorm` is provided with its own unique identifier.

- **Rehabilitation:** Rehabilitations are prescribed to patients for a given amount of time. They are represented by the entity `Rehabilitation`.

- **Diagnosis:** Diagnoses are officially defined and categorized into groups by the *International Classification of Diseases*[1]. In the data model, they are represented by the `Diagnosis` and `Diagnosis Group` entities. Diagnoses are assigned to patients through the relationship `Patient-Diagnosis`.

- **Specification of damage:** Assigns damage, illness, or handicap classifications to patients as keywords, represented by the `Damage Tag` entity.

5.4 REST API

The application layer of the system is represented by a REST API. The implementation of a REST API (abbreviated as "API" in the rest of this chapter) allows the clients to be completely decoupled from the backend, resulting in a cleaner design and codebase. Additionally, the process of updating or implementing a new client application is significantly simplified. The API offers an interface that allows clients to interact with the data in the database through HTTP requests. These requests are handled by endpoints, which can be classified into the following categories:

1. **CRUD operations:** The majority of entities provide five endpoints for CRUD (create, read, update and delete) operations, which include:

   - `GET /entity`: Retrieves all entity records from the database.
   - `GET /entity/{id}`: Retrieves a single record by its identifier.
   - `POST /entity`: Creates a new entity record.
   - `PATCH /entity/{id}`: Updates an existing record.
   - `DELETE /entity/{id}`: Deletes a record.

   Note: The term "entity" in the endpoint paths serves as a placeholder for the true name of the entity.

2. **Authentication:** Includes three endpoints:

[1] https://www.who.int/standards/classifications/classification-of-diseases

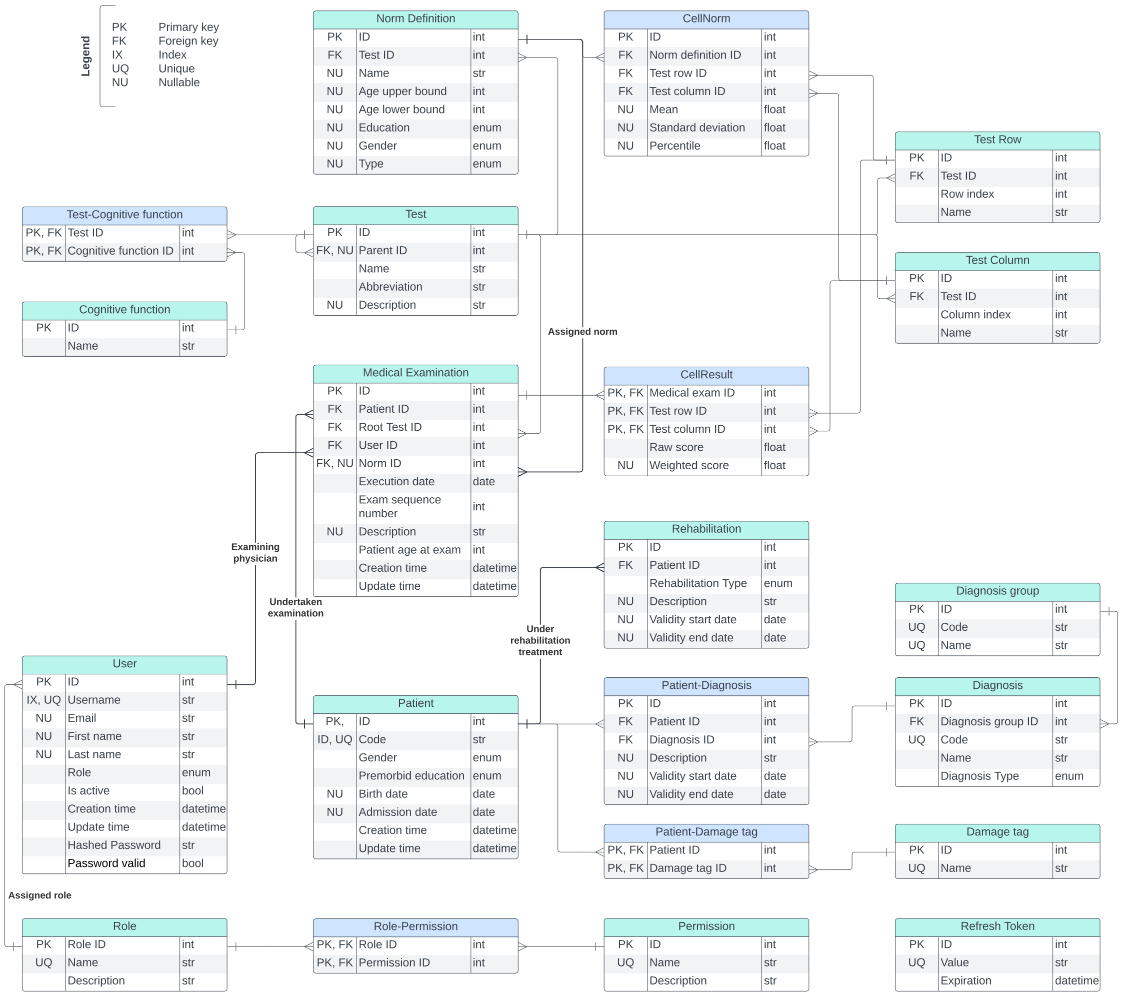


Figure 5.3: The ER diagram of the designed solution.

- `GET /auth/me`: Provides information about the currently logged-in user, including their role and permissions.
- `POST /auth/token`: Returns a short-lived access token based on the provided credentials.
- `POST /auth/refresh`: Returns a new access token for users who are already logged in.

3. **Users:** The management of users is limited to users with the admin role. In addition to the standard CRUD operations, the endpoints provide the following additional functionality:

   - `PUT /users/{id}/force_change_password`: Allows an admin to reset the password of any other user.
   - `PATCH /users/{id}/activation`: Allows an admin to activate or deactivate the account of any user. A user with a deactivated account is restricted from accessing the application.

4. **Visualizations:** These endpoints provide data required for the generation of visualizations:

   - `/analyses/comparison-with-other-patients`:
     Returns both the raw and normed results of the entire reference group, as well as the strict and weak percentile ranks of the patient being compared. The request must specify the compared patient, the reference group and the test of interest.
   - `/analyses/result-history`: Returns the historical record of a patient's results for a specified test.

5. **Exports:** These endpoints allow the users to export data both in `csv` and `xls` formats (see Section 5.7):

   - `/analyses/data-matrix/1/format`: Returns data representing the Data Matrix No. 1.
   - `/analyses/data-matrix/2/format`: Returns data representing the Data Matrix No. 2.

   Note: The term "format" in the endpoint paths serves as a placeholder for format requested.

6. **Exceptions:** Following is a list of endpoints that do not fall in any of the previous categories:

   - `/`: Returns general information about the application, such as the version number, contact details of the author, etc.
   - `/roles`: Returns a list of available user roles.
   - `/diagnoses/assigned`: Returns a list of all diagnoses that are currently assigned to at least one patient.

5.5 Client

The presentation layer of the system is represented by a web application that runs in the web browser. Users at the RCK will be able to use the application without the need to install any additional software, which will facilitate the deployment process. Furthermore, the web-based approach ensures out of the box compatibility across various operating systems.

Access to different views and the corresponding allowed actions is determined by the user's role. The user interface is designed to provide psychologists with simple access to their most frequently performed tasks, such as creating and reviewing examination results. On the main page, clinicians can quickly locate a patient and display their general details, past examinations, and assigned rehabilitations and diagnoses. Figure 5.4 illustrates the links between the application's views. The shallow hierarchy of the application's design increases the efficiency of navigation.

5.6 Data Visualizations

The visualizations incorporated in the system are based on the requirements stated in Section 4.5.1. The final design of the visualizations and the related views was discussed with PhDr. Petra Fiľová, Ph.D. and Mgr. Pavel Viktora, in accordance with the protocol presented in Appendix D.

5.6.1 Comparison with Reference Group

The view that allows clinicians to compare a patient's results with those of an arbitrary reference group includes the following components:

- **Chart:** A histogram of the results of the reference group. The histogram also includes a marker indicating the analyzed patient's results.
- **Filter:** Clinicians can filter the patient list by sex, age, premorbid education, damage specification, and diagnosis.
- **Patient selection:** A list with check boxes allows clinicians to precisely select specific patients that should be included in the visualization.
- **Statistics table:** A table that displays the numerical values of the results of the patient and the related strict and weak percentile ranks.
- **Switch:** Clinicians can switch between raw and normed results.

5.6.2 Historic Development

This view visualizes the historical progression of a patient's test results and is composed of:

- **Chart:** A line chart that plots the patient's historical performance in one or more selected test cells. The chart provides a visual representation of result changes across examinations.
- **Filter:** Allows clinicians to select specific cells of the tests and all their subtests to be included in the visualization.
- **Switch:** A switch between raw and normed results.

5.7 Exports

The system allows to export data in accordance with the requirements specified in Section 4.5.2. The data are exported in the form of two data matrices, which are provided in both `csv` and `xls` formats. The advantage `csv` format is that it is widely supported and can be imported by the majority of data science tools. On the other hand, the `xls` format is a commonly used format among psychologists, and will allow them to further investigate the data in a quick manner. The first data matrix, introduced in Table 5.2, provides users with information about examinations. The second data matrix, introduced in Table 5.3, provides information about patients.

Patient code	**Date**	**Age at the day of examination**	**Subtest 1**	**Subtest 2**	**...**	**Subtest N**
P1	2001-05-01	45	15	24	...	23
P1	2001-06-01	45	17	25		27
P2	2002-12-15	23	14	27		22
...						

Table 5.2: The Data Matrix 1. The data in the matrix are linked to a single specific test. Each record represents a single examination. If a patient was administered the same test multiple times, they will appear in multiple records.

Patient code	**Sex**	**Date of birth**	**Premorbid education**	**Diag. 1**	...	**Diag. N**	**Damage spec. 1**	...	**Damage spec. N**
P1	M	2000-01-01	Primary	YES		NO	NO		NO
P2	F	2000-01-01	Secondary	NO		YES	YES		NO
...									

Table 5.3: The Data Matrix 2. Each record represents a single patient within the database. The record includes general information about the patient, their diagnoses and damage specifications.

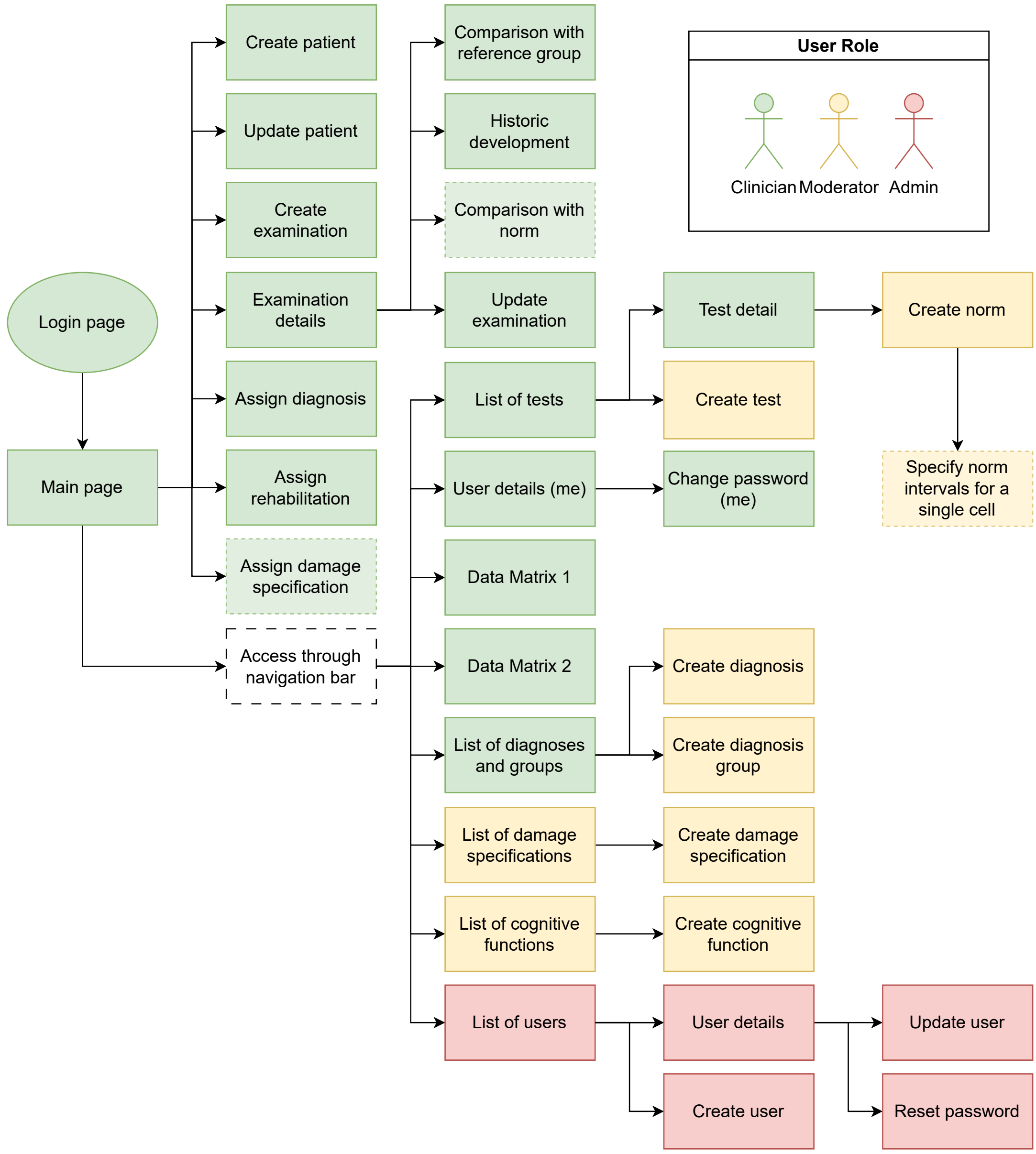


Figure 5.4: All the views implemented in the system and the manner in which users are expected to navigate through them. The elements with dashed borders represent views that are implemented in the form of a dialog.

Chapter 6

Implementation

The following chapter builds upon the design specified in Chapter 5. Section 6.1 introduces the implementation details of the application's backend, with a particular emphasis on the layered architecture of the REST API. The structure of the frontend is presented in Section 6.2, which also details the libraries used and their respective roles within the solution. The chapter concludes with Section 6.3, which discusses the authentication and authorization mechanisms employed by the REST API, based on JWT tokens.

6.1 Backend

The backend of the system consists of two main components: A database server, based on the relational database management system PostgreSQL[1] 14.11, and a REST API server, implemented in Python[2] 3.11 with the use of the FastAPI[3] library. The API implements 76 endpoints, which are described in detail in Section 5.4 and listed in full in Appendix E. An overview of the integration of the API into the overall system architecture is presented in Figure 6.1. The structure of the API is divided into three layers:

1. **Access layer:** Defines the API's communication interface. The package[4] `AL/routers` contains the definitions of the exposed endpoints, including the HTTP methods they use and parameters they require. The `AL/schemes` package defines the structure of the data objects that are used for communicating with the API. The schemes of these objects are implemented using the Pydantic[5] library. This solution allows for the automatic validation of input data and to directly map the data to and from SQLAlchemy[6] models.

2. **Business layer:** This layer acts as the intermediary between the access layer and the data access layer. It implements the core business logic of the application. The `BL/services/` package contains modules that manage operations on database entities. In order to communicate with the database, the modules rely on a database session that gets injected from the data access layer.

[1] https://www.postgresql.org/

[2] https://www.python.org/

[3] https://fastapi.tiangolo.com/

[4] The terms *package* and *module* refer to folders and files, respectively. This is a standard naming convention in Python.

[5] https://docs.pydantic.dev

[6] https://www.sqlalchemy.org/

3. **Data access layer:** The layer builds on the ORM[7] library SQLAlchemy, which abstracts database entities into Python models. The model of the database is defined in the module `DAL/models.py` and follows the schema defined in 5.3. The model fully defines cascading deletions of child entities. The module `DAL/database.py` defines the parameters of the database session, which is used by the application to connect and communicate with the database.

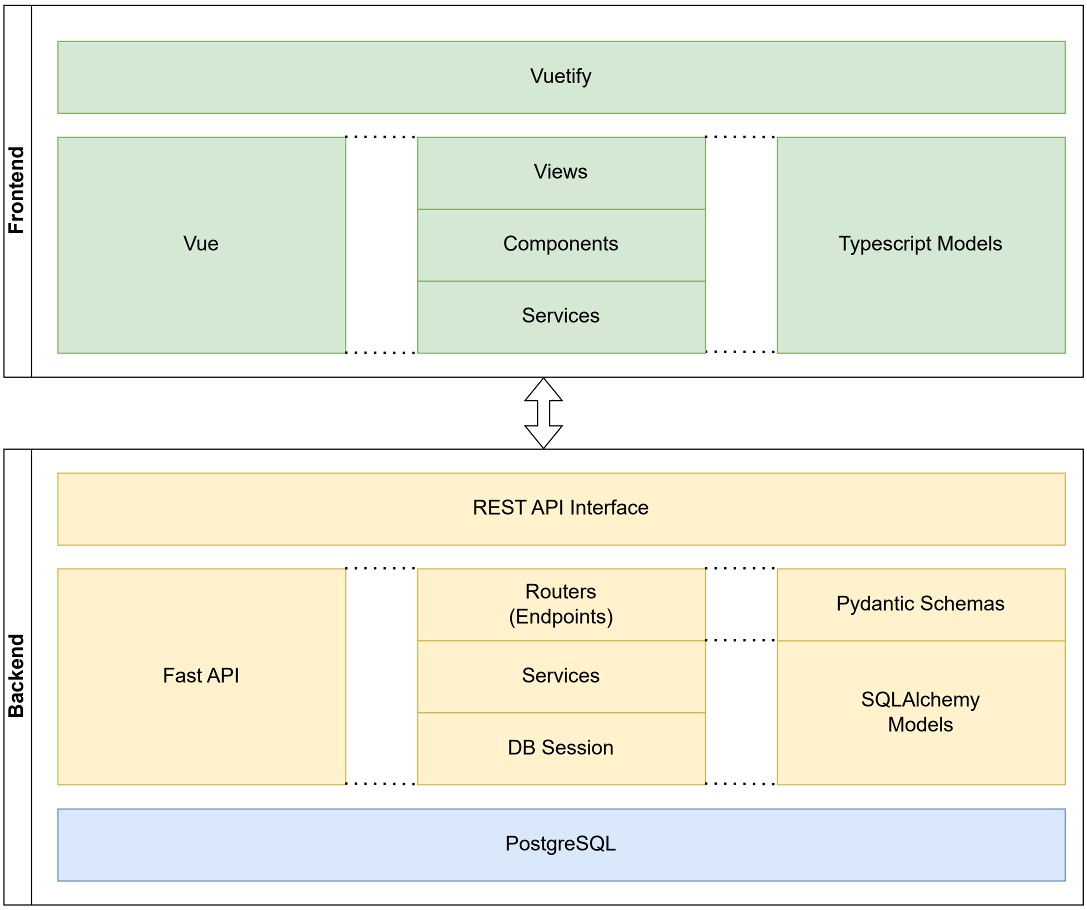


Figure 6.1: The architecture of the implemented system. The frontend communicates with the backend through a REST API interface.

6.2 Frontend

In order to achieve a high degree of speed and interactivity, the frontend of the application is implemented as a single-page web application. The solution is based on the Vue.js[8] 3 framework and the TypeScript[9] programming language. The layout adopts a coherent visual style that adheres to the principles of Material Design[10]. The core views are presented

[7]Object-relational mapping
[8]https://vuejs.org/
[9]https://www.typescriptlang.org/
[10]https://material.io/

in Appendix A. The code follows the Composition API style, which enables the creation of clean, efficient logic that can be reused across multiple components. Additionally, the Composition API allows to group together code related to the same logical concern [40]. The structure of the code of the solution can be divided into three main concepts, each represented by a corresponding folder:

- `views/`: The views are Vue components designed to render entire pages. They serve as the root nodes in the component hierarchy tree used within the application. Figure 6.2 illustrates an example of the hierarchical structure of these components.
- `components/`: This folder contains Vue components intended for reuse throughout the application. The term "components" is used here in alignment with the Vue framework's terminology.
- `client/`: The client is generated using the open-source CLI tool *Openapi Typescript Codegen*[11] and is based on the API specification introduced in Section 5.4. This generated client simplifies the process of calling API endpoints from within the Vue application. It includes three key components essential for API interaction. Firstly, it contains models that define the data structures utilized for communication with the API. Secondly, it provides services that define the requests available to the application. Lastly, it implements core functionality responsible for managing HTTP communication with the API.

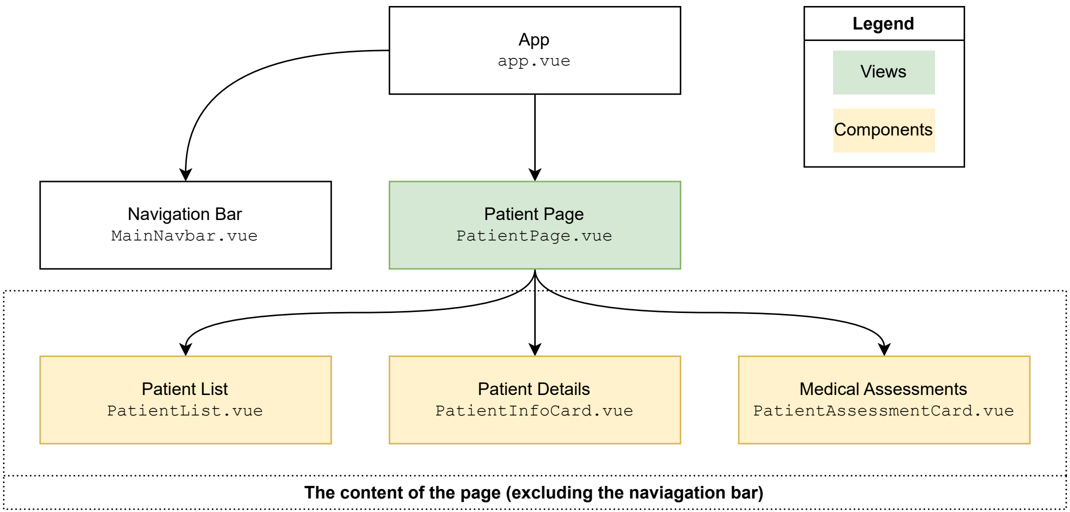


Figure 6.2: The hierarchical structure of Vue components on the main page (see Figure A.1) of the application. The `PatientPage.vue` file is located in the `views/` folder, while the remaining components highlighted in yellow are located in the folder `components/`. Note: Although all elements in this diagram are technically Vue components, the diagram distinguishes between *views* and *components* in the way they are used within the context of the application.

The application is powered by multiple libraries that enhance the application's functionality. Following is a list of integrated libraries and their role in the application:

[11] https://github.com/ferdikoomen/openapi-typescript-codegen

- **Vue Router**[12]: The integration of Vue Router allows users to navigate between different views within the application without the need to reload the page from the server [39]. Routes and their associated URLs are configured in the `plugins/router.ts` file. The router takes care of several critical functions. It verifies user permissions and redirects users from restricted or non-existent pages. Moreover, it can automatically log in users if a valid refresh token is present in the client's storage.

- **Vuetify**[13]: The user interface of the system is built using the Vuetify library that offers a wide range of components, such as buttons, lists, tables and other elements, which follow the material design style. Additionally, the library provides methods for creating the layout.

- **Apache Echarts**[14]: The library provides visualization tools in the form of various types of charts. The application visualizes psychological data using line and bar charts provided by this library. The visualizations are implemented in accordance with the design defined in Section 5.6.

- **Pinia**[15]: The Pinia library is used for the state management of the application. The application makes use of Pinia storages primarily to share data between components and avoid the use of the traditional `props` and `emits` pattern. Notably, the `stores/userStore.ts` and `stores/tokenStore.ts` play a crucial role by managing the information of the currently logged-in user and the JWT tokens, respectively.

- **CASL**[16]: The application uses the CASL library to conditionally render UI elements, ensuring that the visibility and interactivity of these elements align with the permissions of the currently logged-in user.

6.3 Authentication and Authorization

The application implements an authentication and authorization mechanism that utilizes access and refresh JWT[17] tokens. The complete authentication flow is illustrated in Figure 6.3. The REST API provides two endpoints for the purpose of user authentication:

- `POST /auth/token`: To log in, clients must submit their credentials to this endpoint using the content type `application/x-www-form-urlencoded`. If successful, the endpoint returns a short-lived access token that authorizes the user to perform requests on the API according to their role. The endpoint requires valid credentials to be sent with the request. The response also includes a long-lived refresh token, which is utilized in the second endpoint.

- `POST /auth/refresh`: This endpoint allows clients to obtain a new access token using a refresh token. This eliminates the necessity for clients to re-authenticate with credentials each time their access token expires. The endpoint returns a new pair of access and refresh tokens. Since FastAPI does not provide the refresh token

[12] https://router.vuejs.org/
[13] https://vuetifyjs.com
[14] https://echarts.apache.org
[15] https://pinia.vuejs.org/
[16] https://casl.js.org
[17] JSON Web Token: https://jwt.io/

functionality by default, the API implements a custom solution that stores the active refresh tokens into the database manually.

Authorization for REST API requests relies on valid access tokens. Endpoints require that the request includes the `Authorization` header containing the access token in the following format:

```
Authorization: Bearer <token>
```

This header must be included in the HTTP request sent to the server. If the token is valid, the server will satisfy the request. In the opposite case, the server will respond with a 401 Unauthorized HTTP error. When the short-lived access token expires, the client can obtain a new one using the refresh token.

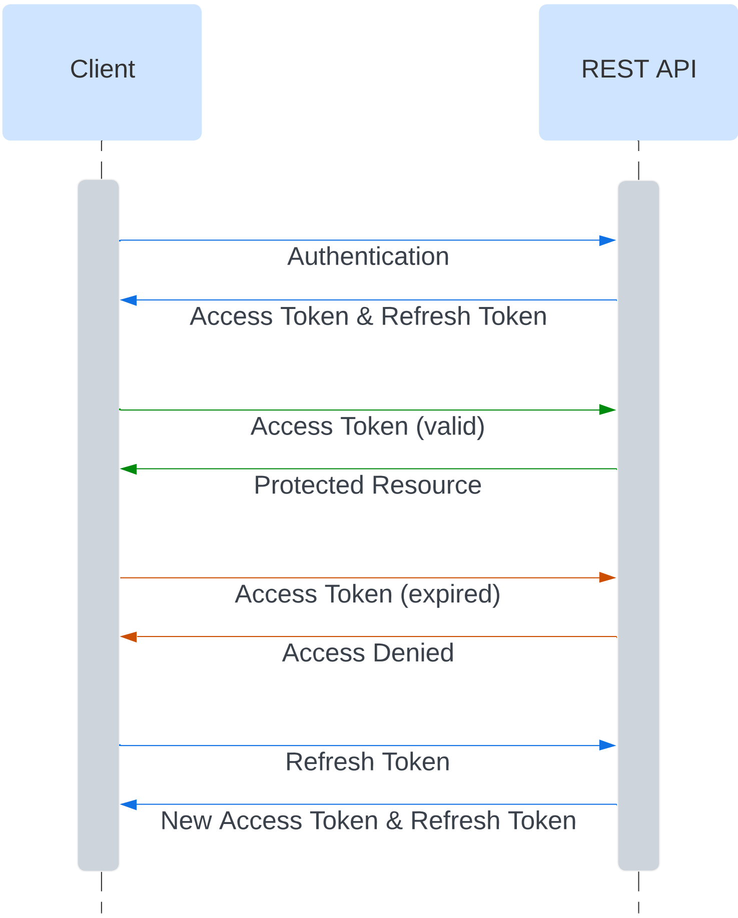


Figure 6.3: A sequence diagram of the JWT authentication and authorization process. The diagram uses a color scheme to differentiate between the types of requests. Blue represents authentication, green indicates successful authorization, and red unsuccessful authorization.[18]

[18]The figure is inspired by the article: https://medium.com/@chauhanshubham19765/jwt-refresh-token-61823e888bc7

Chapter 7

Testing

This chapter presents an overview of the application's testing methodology. Initially, unit testing was conducted to verify the correct functionality of the REST API, as detailed in Section 7.1. As development progressed, user acceptance testing was implemented during the final phases to ensure that the application aligned with the expectations of the Rehabilitation Center Kladruby. The outcomes of the user acceptance testing are discussed in Section 7.2.

7.1 Unit Testing

During the early stages of development of the REST API, a set of unit tests was implemented to verify the functionality of the endpoints. The intention was to test each individual endpoint in an isolated manner across a variety of test scenarios. The tests were implemented using the Pytest[1] library. The testing process utilized a pre-seeded database whose state was reset before every test run. The test cases for most entities included the following scenarios:

- Standard CRUD operations.
- Retrieval of all records of an entity.
- Creation of entities with invalid data.
- Creation of duplicate entities.
- Testing of filters (if supported by the endpoint).

The test cases examined both the returned data and the response code. The request bodies sent to the REST API were defined as pytest fixtures. Figure 7.1 demonstrates an example of a single unit test.

[1] https://docs.pytest.org

```python
@pytest.fixture
def user1() ->UserPostRequest:
    return UserPostRequest(
        username="tom",
        first_name="Tom",
        # ...
    )

def test_create_user_duplicate(user1):
    """Test that a user cannot be created twice."""
    response1 =client.post("/users/", json=user1.model_dump())
    assert response1.status_code ==201

    response2 =client.post("/users/", json=user1.model_dump())
    assert response2.status_code ==409
```

Figure 7.1: An example of a definition of a unit test used in the early process of testing. The fixture `user1` is automatically injected into the test case `test_create_user_duplicate` and is utilized as the body of the POST request.

Given the dynamic nature of the development process and the frequent changes to the data model, unit tests were eventually abandoned and disregarded as overly time-consuming during the later stages of development.

7.2 User Acceptance Testing

The user acceptance testing phase was conducted in April 2024, when a nearly finalized version of the system was presented to the representatives of the Rehabilitation Center Kladruby (RCK). The objective of this stage was to validate the existing application against the requirements of the RCK. The overall feedback was predominantly positive, resulting in the following enhancements being introduced into the system:

- Added filters and detailed patient selection to the visualization comparison that compares a patient with a reference group.
- Added the possibility to include multiple subtests in a single chart to the visualization of historical results.
- Introduced the possibility to automatically load diagnoses and diagnosis groups from the website of the *General Health Insurance Company of the Czech Republic*[2]. The data are available in the form of a `csv` file that can be loaded into the database using the standalone module `load_diagnoses.py`.
- Made minor modifications to the texts displayed in the user interface.
- Enabled rehabilitations and patient diagnoses to be recorded without specifying their start or end dates.
- Included information about the examining psychologist in the examination record.
- Removed the direct assignment of a psychologist to a patient.

[2] https://www.vzp.cz/poskytovatele/ciselniky/mezinarodni-klasifikace-nemoci

In addition to these successfully implemented changes, two additional features were proposed but have not been incorporated. The representatives proposed to add an alert in the user interface that notifies a user when an examination result entered is more than three standard deviations from the norm's mean. The second suggestion was to integrate support for medications that would be automatically fetched from a public API. However, due to the non-trivial nature of these changes and time constraints, the implementation has been postponed for future updates.

Another subject of discussion was the future of the project, including its deployment, maintenance process, and future updates. It is expected that the project will be deployed at the RCK during the summer of 2024. Furthermore, Mgr. Pavel Viktora suggested that the system might enhance the process of results comparison by adding a possibility to define thresholds for identifying scores that are above and below average.

In the final communication via email, PhDr. Petra Fiľová, Ph.D. expressed satisfaction with the system's design, noting its aesthetic appeal and user-friendly interface. She reported a smooth experience in navigating the application and specifically highlighted the importance of the inclusion of data matrices, which she acknowledged as valuable for future statistical analysis. She confirmed that all previously discussed points were addressed to her satisfaction.

Chapter 8

Conclusion

The objective of this thesis was to design and develop an information system in collaboration with the Rehabilitation Center Kladruby to manage and visualize psychological diagnostic data. The system aims to achieve two primary improvements: firstly, to assist clinicians at the rehabilitation center in the evaluation of psychological test results, and secondly, to establish a database that could serve as a valuable resource for the creation of new psychometric assessment norms.

In order to design a meaningful solution, it was necessary to gain an understanding of the principles of testing in psychology, including the underlying methods of descriptive statistics and the process of test evaluation based on psychometric norms. Moreover, the theoretical background of the processes involved in the design of information systems with a focus on data visualizations was studied.

The specifications of the system were discussed with the representatives of the Rehabilitation Center Kladruby on multiple occasions. Based on these specifications, an information system comprising three layers was designed. The system is divided into a database, an application server, and a client. This design allows for the development of multiple independent clients without the necessity of modifying the server, a feature that may prove advantageous in the future. The application was implemented using a combination of FastAPI and Vue.js.

The final application meets the majority of the requirements and has successfully passed user acceptance testing. Clinicians at the Rehabilitation Center Kladruby will have the possibility to digitalize the process of collecting and evaluating psychological tests, which is expected to reduce their workload and enhance the efficiency of patient treatment. The results were presented at the Excel@FIT 2024[1] student conference, offering a chance for a wider audience to engage with the application and provide valuable feedback [13].

However, certain features were not implemented due to time constraints and must be implemented before the final deployment. These include user interface responsiveness and client scalability, particularly in the phase of loading data from the application server. Other potential future improvements include support for medications and different types of norms. From a long-term perspective, if the system proves successful at the Rehabilitation Center Kladruby, it may be offered to other similar facilities in the future.

[1] https://excel.fit.vutbr.cz/

Bibliography

[1] Aeberhard, S. and Forina, M. *Wine* [online]. UCI Machine Learning Repository, 1991. DOI: 10.24432/C5PC7J. Available at: https://doi.org/10.24432/C5PC7J.

[2] American Educational Research Association, American Psychological Association and National Council on Measurement in Education. *Standards for educational and psychological testing.* American Educational Research Association, 2014. ISBN 978-0-935302-35-6.

[3] Andrade, C. Z scores, standard scores, and composite test scores explained. *Indian Journal of Psychological Medicine.* Oct 2021, vol. 43, no. 6, p. 555–557. DOI: 10.1177/02537176211046525. Available at: https://journals.sagepub.com/doi/10.1177/02537176211046525.

[4] Andrienko, N., Andrienko, G., Fuchs, G., Slingsby, A., Turkay, C. et al. *Visual Analytics for Data Scientists.* Springer, 2020. ISBN 978-3-030-56145-1.

[5] Anscombe, F. J. Graphs in Statistical Analysis. *The American Statistician.* [American Statistical Association, Taylor & Francis, Ltd.]. 1973, vol. 27, no. 1, p. 17–21. ISSN 00031305. Available at: http://www.jstor.org/stable/2682899.

[6] Arnheim, R. *Visual Thinking.* University of California Press, 1972. ISBN 978-0520018716.

[7] Arnheim, R. *Art and Visual Perception: A Psychology of the Creative Eye.* University of California Press, 1974. ISBN 0-520-24383-8.

[8] Bernus, P., Mertins, K. and Schmidt, G. *Handbook on Architectures of Information Systems.* 2nd ed. Berlin, Heidelberg, Germany: Springer-Verlag, 2006. ISBN 978-3-540-25472-0.

[9] Beynon Davies, P. *Significance: exploring the nature of information, systems and technology.* Basingstoke, England: Palgrave Macmillan, 2011. ISBN 978-0-230-27519-5.

[10] Bradley, J. The insidious L-shaped distribution. *Bulletin of the Psychonomic Society.* Aug 1982, vol. 20, no. 2, p. 85–88. DOI: 10.3758/bf03330089. Available at: https://link.springer.com/article/10.3758/BF03330089.

[11] Britannica, The Editors of Encyclopaedia. *Projective test* [online]. Feb 2020. Encyclopedia Britannica. Available at: https://www.britannica.com/science/projective-test.

[12] Chaffey, D. and Wood, S. *Business Information Management: Improving Performance Using Information Systems*. Harlow, England: Pearson Education, 2005. ISBN 978-0-273-68655-2.

[13] Chimenti, A. *Processing and Visualization of Psychological Diagnostic Data*. 2024. Available at: https://excel.fit.vutbr.cz/submissions/2024/068/68.pdf.

[14] Cronbach, L. J. *Essentials of Psychological Testing*. Harper & Brothers, 1949.

[15] DeVellis, R. F. Classical Test Theory. *Medical Care*. Lippincott Williams & Wilkins. 2006, vol. 44, no. 11, p. S50–S59. ISSN 00257079. Available at: http://www.jstor.org/stable/41219505.

[16] Furr, R. M. *Psychometrics: An introduction*. 4th ed. Sage, 2022. ISBN 978-1-0718-2407-8.

[17] García, G. *Data Visualization and Analysis in Second Language Research: Second language acquisition research series*. Routledge, 2021. ISBN 978-0-367-46961-0.

[18] Goodfellow, I., Bengio, Y. and Courville, A. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[19] Hogan, T. P. *Psychological Testing: A Practical Introduction*. 3rdth ed. Hoboken, NJ: John Wiley & Sons, 2015. ISBN 978-1-118-55412-8.

[20] Hunt, M. *Dějiny psychologie*. Portál, 2006. ISBN 80-7367-175-1.

[21] Jackson, S. L. *Research Methods and Statistics: A Critical Thinking Approach*. 5th ed. Boston: Cengage Learning, 2016. ISBN 978-1-305-25779-5.

[22] Knaflic, C. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley, 2015. ISBN 978-1-119-00225-3.

[23] Laudon, K. and Laudon, J. *Management information systems: Managing the Digital Firm*. 16th ed. Harlow, England: Pearson, 2020. ISBN 978-1-292-29656-2.

[24] Marakas, G. M. and O'Brien, J. A. *Introduction to information systems*. 16th ed. New York, NY: McGraw-Hill/Irwin, 2013. ISBN 978-0-07-337688-2.

[25] Myatt, G. J. *Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining*. Hoboken, N.J: John Wiley & Sons, Inc., 2007. ISBN 978-0-470-07471-8.

[26] Playfair, W. *The Commercial and Political Atlas: Representing, by Means of Stained Copper-plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England During the Whole of the Eighteenth Century*. 3rd ed. T. Burton, 1801.

[27] Reese, G. *Database Programming with JDBC & Java*. 2nd ed. Sebastopol, CA: O'Reilly Media, Inc., 2000. ISBN 9781565926165.

[28] Rehabilitační ústav Kladruby. *Výroční zpráva 2022* [online]. 2022 [cit. Apr 20, 2024]. Available at: https://www.rehabilitace.cz/o-nas/vyrocni-zpravy/.

[29] Rehabilitační ústav Kladruby. *Historie* [online]. 2024 [cit. Apr 20, 2024]. Oficiální stránky Rehabilitační ústav Kladruby. Available at: https://www.rehabilitace.cz/o-nas/historie/.

[30] Rehabilitační ústav Kladruby. *O nás: Naše motto, vize, poslání a hodnoty* [online]. 2024 [cit. Jan 24, 2024]. Oficiální stránky Rehabilitační ústav Kladruby. Available at: https://www.rehabilitace.cz/o-nas/nase-motto-vize-poslani-a-hodnoty/.

[31] Rowley, J. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*. 2007, vol. 33, no. 2, p. 163–180. DOI: 10.1177/0165551506070706. Available at: https://doi.org/10.1177/0165551506070706.

[32] Saraswathi, R. *Four Architecture Choices for Application Development in the Digital Age* [online]. 2020 [cit. Jan 20, 2024]. Available at: https://www.ibm.com/blog/four-architecture-choices-for-application-development/.

[33] Schwabish, J. *Better Data Visualizations: A guide for scholars, researchers, and wonks*. Columbia University Press, 2021. ISBN 978-0-231-19311-5.

[34] Seashore, H. G. Methods of Expressing Test Scores. *Test Service Bulletin*. Jan 1955, vol. 48, p. 7–10. Available at: https://eric.ed.gov/?id=ED079347.

[35] Skiena, S. S. *The Data Science Design Manual*. Cham, Switzerland: Springer, 2017. ISBN 978-3-319-55443-3.

[36] Tufte, E. *The Visual Display of Quantitative Information*. 2nd ed. Graphics Press, 2001. ISBN 978-1930824133.

[37] Urbánek, T. Stav české psychologické diagnostiky a evropský model recenze testu. *TESTFÓRUM*. May 2010, vol. 1, no. 1, p. 2–5. DOI: 10.5817/tf2010-1-1. Available at: https://testforum.cz/article/view/TF2010-1-1.

[38] Urbánek, T., Denglerová, D. and Širůček, J. *Psychometrika: měření v psychologii*. 1st ed. Praha: Portál, 2011. ISBN 978-80-7367-836-4.

[39] Vue Router. *Getting Started* [online]. [cit. May 4, 2024]. Available at: https://router.vuejs.org/guide/.

[40] Vue.js. *Composition API FAQ* [online]. [cit. May 3, 2024]. Available at: https://vuejs.org/guide/extras/composition-api-faq.html.

[41] WebMD, Editorial Contributors. *Neuropsychological Tests* [online]. 2022. WebMD. Available at: https://www.webmd.com/brain/neuropsychological-test.

[42] Wilke, C. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. O'Reilly Media, 2019. ISBN 978-1-492-03108-6. Available at: https://clauswilke.com/dataviz.

[43] Zachry, M. and Thralls, C. An Interview with Edward R. Tufte. *Technical Communication Quarterly*. Routledge. 2004, vol. 13, no. 4, p. 447–462. DOI: 10.1207/s15427625tcq1304_5. Available at: https://doi.org/10.1207/s15427625tcq1304_5.

Appendix A

User Interface of the Final Application

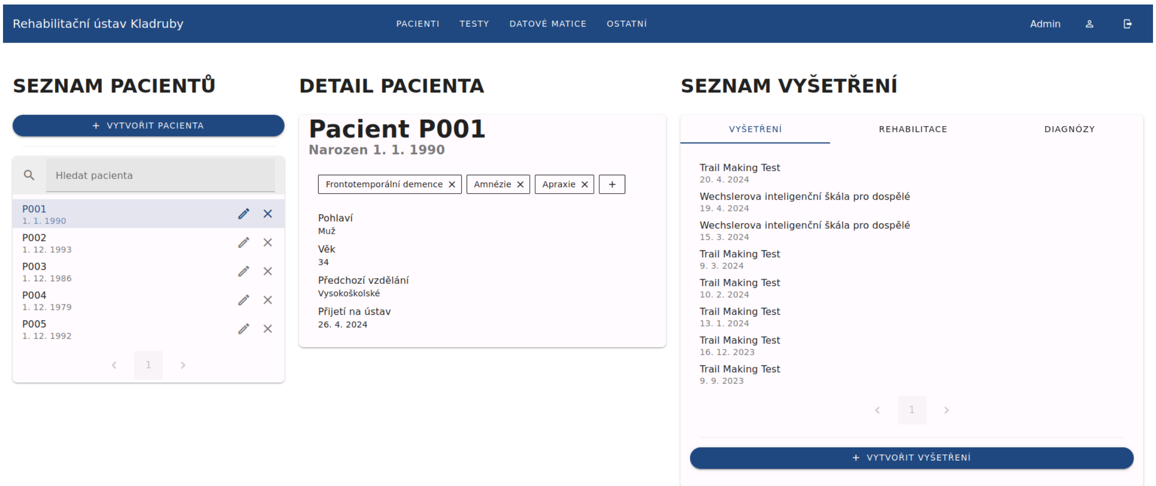


Figure A.1: Main page.

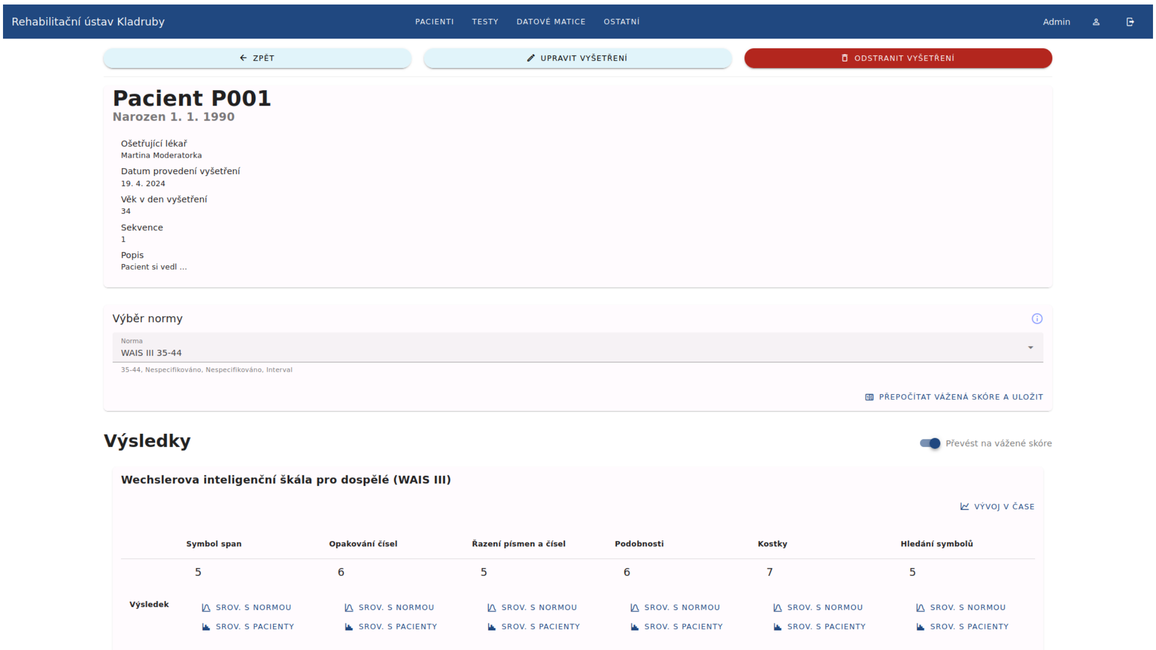


Figure A.2: Examination result page.

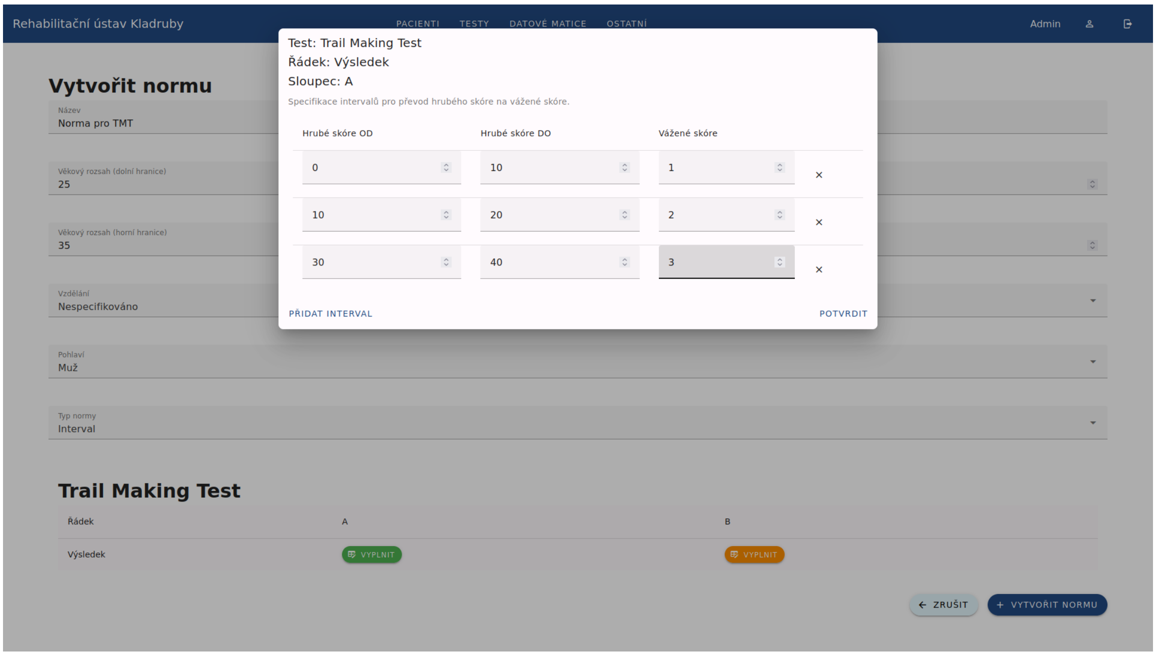


Figure A.3: Norm specification using intervals.

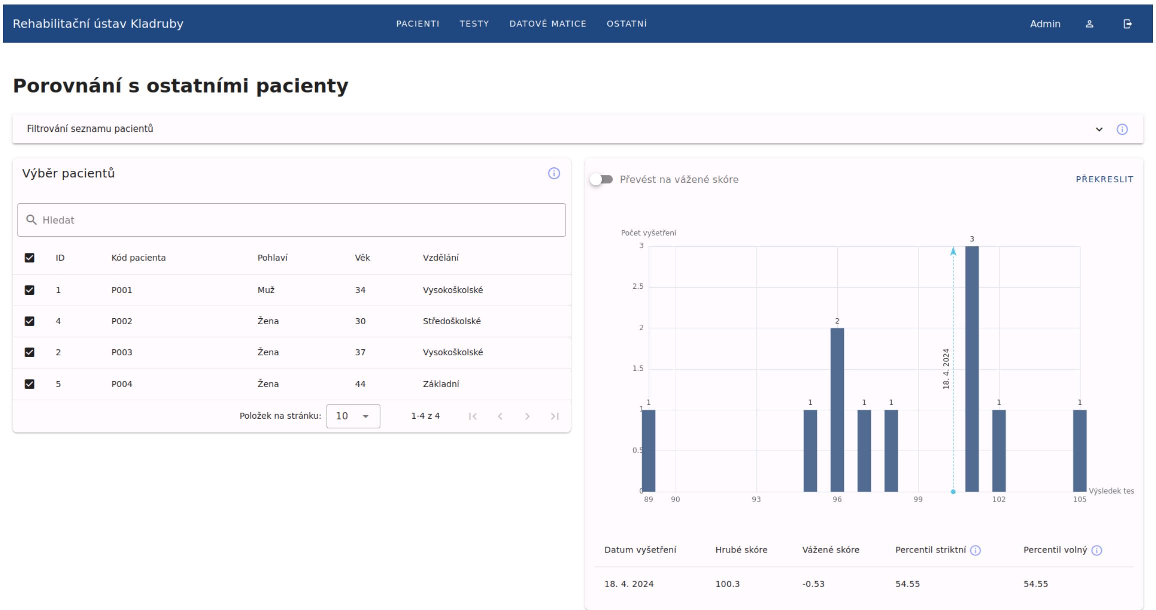


Figure A.4: Visualization of comparison with reference group.

Figure A.5: Visualization of historical results.

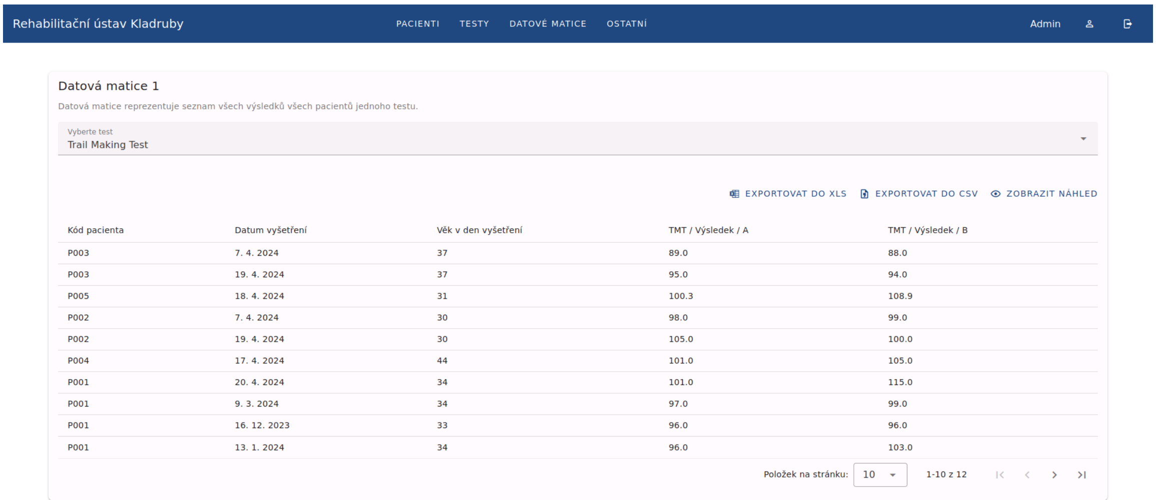


Figure A.6: Data Matrix 1.

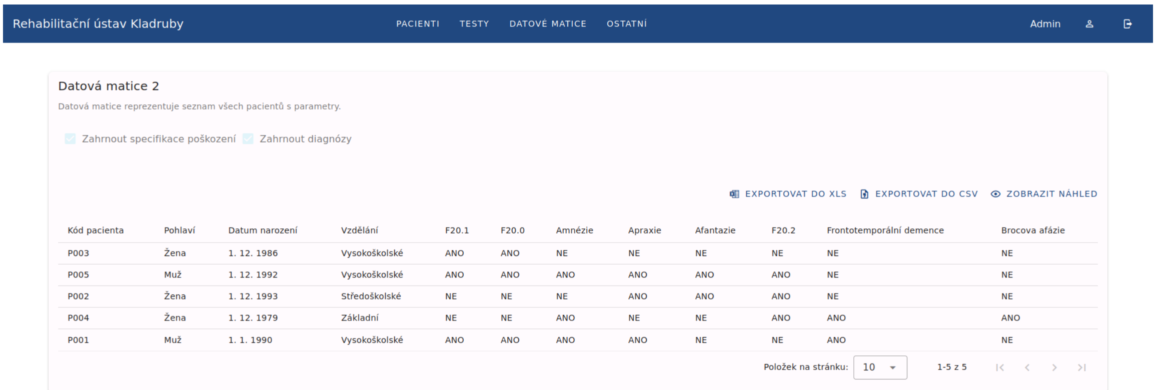


Figure A.7: Data Matrix 2.

Appendix B

Prototype

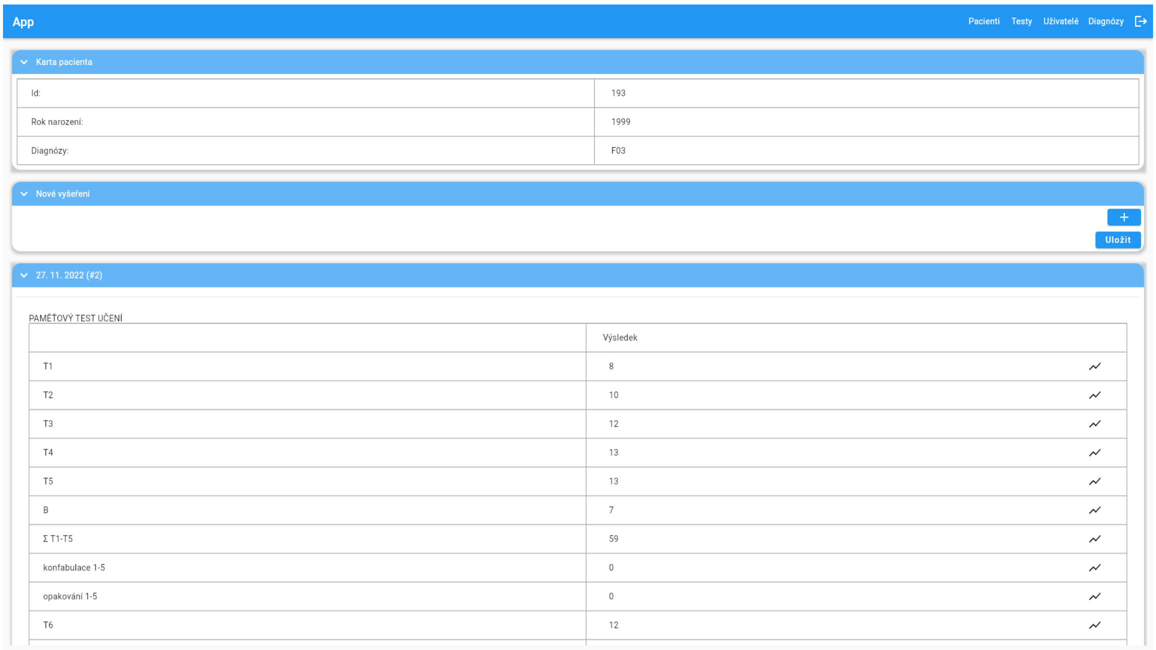


Figure B.1: Patient detail page with examination results.

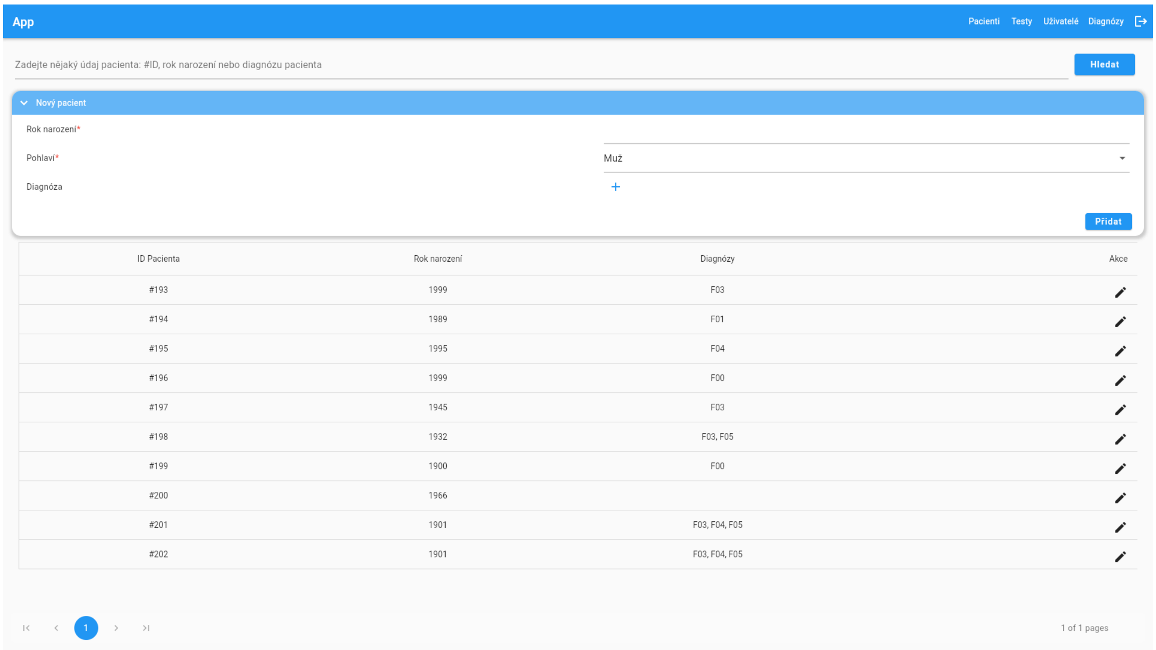


Figure B.2: List of patients.

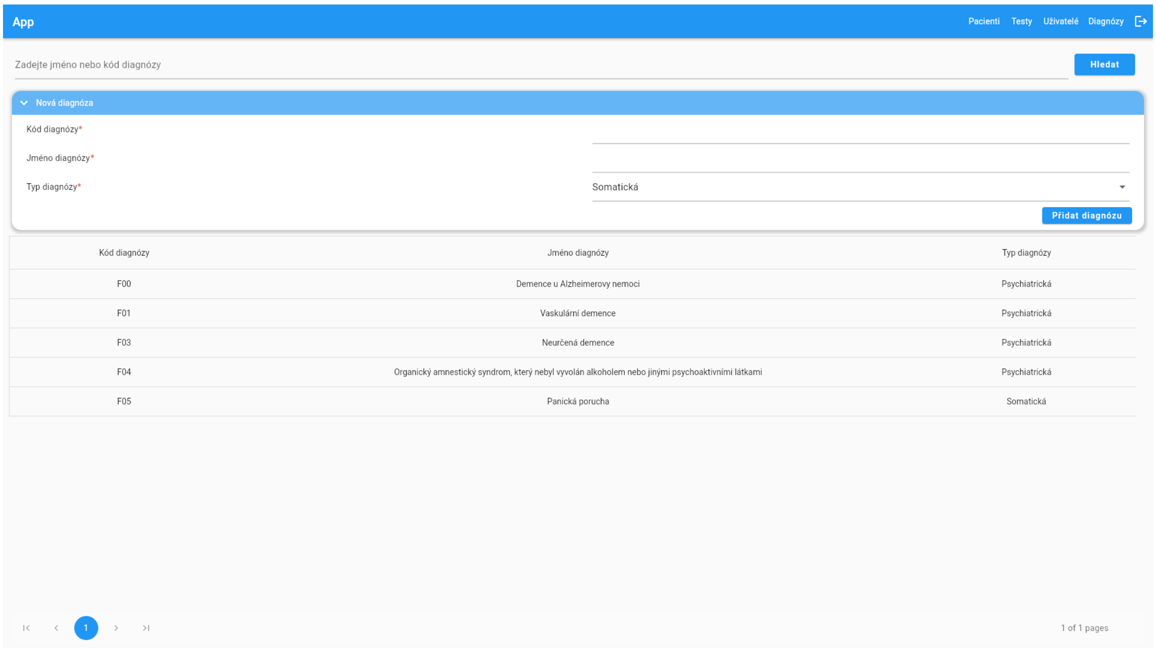


Figure B.3: List of diagnoses.

Appendix C

Tests Used at the Rehabilitation Center Kladruby

Some of the tests used at the Rehabilitation Center Kladruby. The list is not exhaustive:

- Auditory Verbal Learning Test
- Wechsler Memory Scale
- Wechsler Adult Intelligence Scale III
- Rey–Osterrieth Complex Figure
- Trail Making Test
- Stroop Color and Word Test 1-2-3
- Benton Visual Retention Test
- Visual Memory Test
- Boston Naming Test 15
- Verbal Fluency test
- Bells Test

Appendix D

Visualization Design Protocols

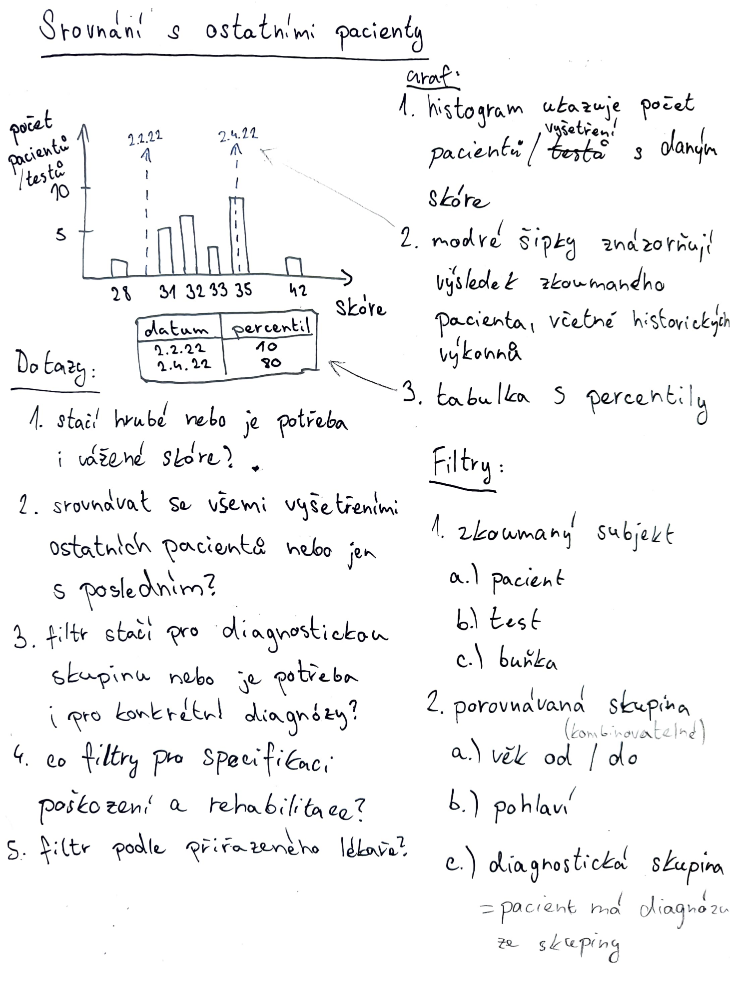


Figure D.1: The protocol was used to analyze the specifications of the visualization used to compare the patient's results with a reference group.

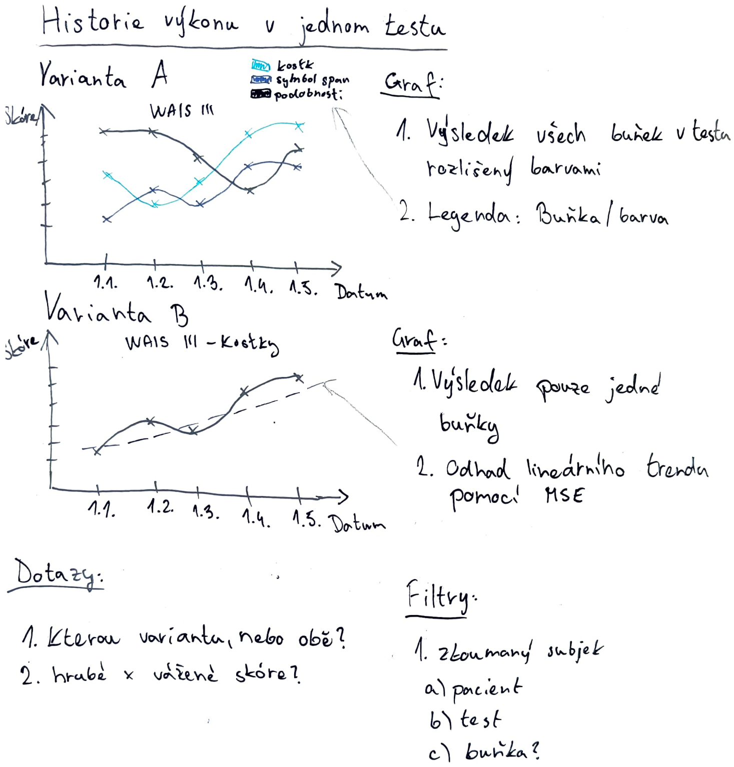


Figure D.2: The protocol was used to analyze the specifications of the visualization used to track the patient's progress over time.

Appendix E

REST API Endpoints

- `GET /auth/me` - Retrieve User Info
- `POST /auth/token` - Login And Get Token
- `POST /auth/refresh` - Refresh Access Token
- `GET /users/` - Get All Users
- `POST /users/` - Create User
- `GET /users/{id}` - Get User By Id
- `DELETE /users/{id}` - Delete User
- `PATCH /users/{id}/details` - Update User Details
- `PATCH /users/{id}/password` - Update Password
- `PUT /users/{id}/force_change_password` - Force Change Password
- `PATCH /users/{id}/activation` - Change Activation
- `GET /patients/` - Get All Patients
- `POST /patients/` - Create Patient
- `GET /patients/{id}` - Get Patient By Id
- `PATCH /patients/{id}` - Update Patient Details
- `DELETE /patients/{id}` - Delete Patient
- `GET /patients/{id}/rehabilitations` - Get All Rehabilitations For Patient
- `POST /patients/{id}/rehabilitations` - Create Rehabilitation For Patient
- `GET /patients/rehabilitations/{rehab_id}` - Get Rehabilitation For Patient By Id
- `PATCH /patients/rehabilitations/{rehab_id}` - Update Rehabilitation For Patient

- `DELETE /patients/rehabilitations/{rehab_id}` - Delete Rehabilitation For Patient
- `GET /patients/{id}/diagnoses` - Get All Diagnoses For Patient
- `POST /patients/{id}/diagnoses` - Create Diagnosis For Patient
- `GET /patients/{id}/diagnosis_groups` - Get All Diagnosis Groups For Patient
- `GET /patients/diagnoses/{patient_diagnosis_id}` - Get Diagnosis For Patient
- `PATCH /patients/{id}/diagnoses/` - Update Diagnosis For Patient
- `DELETE /patients/{id}/diagnoses/{diagnosis_id}` - Delete Diagnosis For Patient
- `GET /patients/{id}/damage_tags` - Get All Damage Tags For Patient
- `POST /patients/{id}/damage_tags/{damage_tag_id}` - Add Damage Tag For Patient
- `DELETE /patients/{id}/damage_tags/{damage_tag_id}` - Remove Damage Tag For Patient
- `GET /patients/{id}/examinations` - Get All Examinations For Patient
- `POST /patients/{id}/examinations` - Create Examination For Patient
- `GET /patients/examinations/{exam_id}` - Get Examination By Id
- `PATCH /patients/examinations/{exam_id}` - Update Examination For Patient
- `DELETE /patients/examinations/{exam_id}` - Delete Examination For Patient
- `PATCH /patients/examinations/{exam_id}/norm/{norm_id}` - Change Examination Norm
- `GET /diagnoses/` - Get All Diagnoses
- `POST /diagnoses/` - Create Diagnosis
- `GET /diagnoses/assigned` - Get All Assigned Diagnoses
- `GET /diagnoses/{id}` - Get Diagnosis By Id
- `PATCH /diagnoses/{id}` - Update Diagnosis
- `DELETE /diagnoses/{id}` - Delete Diagnosis
- `GET /diagnosis_groups/` - Get All Diagnosis Groups
- `POST /diagnosis_groups/` - Create Diagnosis Group
- `GET /diagnosis_groups/{id}` - Get Diagnosis Group By Id
- `PATCH /diagnosis_groups/{id}` - Update Diagnosis Group

- `DELETE /diagnosis_groups/{id}` - Delete Diagnosis Group
- `GET /cognitive_functions/` - Get All Cognitive Functions
- `POST /cognitive_functions/` - Create Cognitive Function
- `GET /cognitive_functions/{id}` - Get Cognitive Function By Id
- `PATCH /cognitive_functions/{id}` - Update Cognitive Function
- `DELETE /cognitive_functions/{id}` - Delete Cognitive Function
- `GET /damage_tags/` - Get All Damage Tags
- `POST /damage_tags/` - Create Damage Tag
- `GET /damage_tags/{id}` - Get Damage Tag By Id
- `PATCH /damage_tags/{id}` - Update Damage Tag
- `DELETE /damage_tags/{id}` - Delete Damage Tag
- `GET /tests/` - Get All Root Tests
- `POST /tests/` - Create Test
- `GET /tests/{id}` - Get Test By Id
- `PATCH /tests/{id}/details` - Update Test Details
- `DELETE /tests/{id}/root` - Delete Test Root
- `GET /tests/{id}/norms` - Get Test Norms
- `POST /tests/{id}/norms` - Create Test Norm
- `GET /tests/{id}/norms-full` - Get Test Norms
- `GET /tests/norms/{norm_id}` - Get Test Norm By Id
- `PATCH /tests/norms/{norm_id}` - Update Test Norm
- `DELETE /tests/norms/{norm_id}` - Delete Test Norm
- `GET /roles/` - Get All Roles
- `POST /analyses/comparison-with-other-patients` - Get Comparison With Other Patients
- `POST /analyses/result-history` - Get Result History
- `POST /analyses/data-matrix/1/csv` - Get Data Matrix 1 Csv
- `POST /analyses/data-matrix/1/xls` - Get Data Matrix 1 Xls
- `POST /analyses/data-matrix/2/csv` - Get Data Matrix 2 Csv
- `POST /analyses/data-matrix/2/xls` - Get Data Matrix 2 Xls
- `GET /` - Root