The main deliverable of the team assignment is a 5-page project report (excl. executive summary, figures, references and appendices) which is to be submitted as a .pdf file via ILIAS no later than the deadline 21.07 12:00. In addition to the report we also expect you to upload your Python code in the form of an annotated Jupyter notebook (.ipynb)

Report details the team project, from the business problem through the data science problem and solution, to recommendations and practical relevance.

As presentation is a key component of a successful data science project we will consider it in our evaluation.

The report should be written clearly and professionally and include the following sections:

1. Cover page with informative title, team number and member names

2. One-page executive summary: summarizes the entire report for a non-technical manager (the business problem, data, the analytics solution, implications and recommendations)

3. Detailed report:

(a) Problem description (business goal and data science goal)

(b) Data description

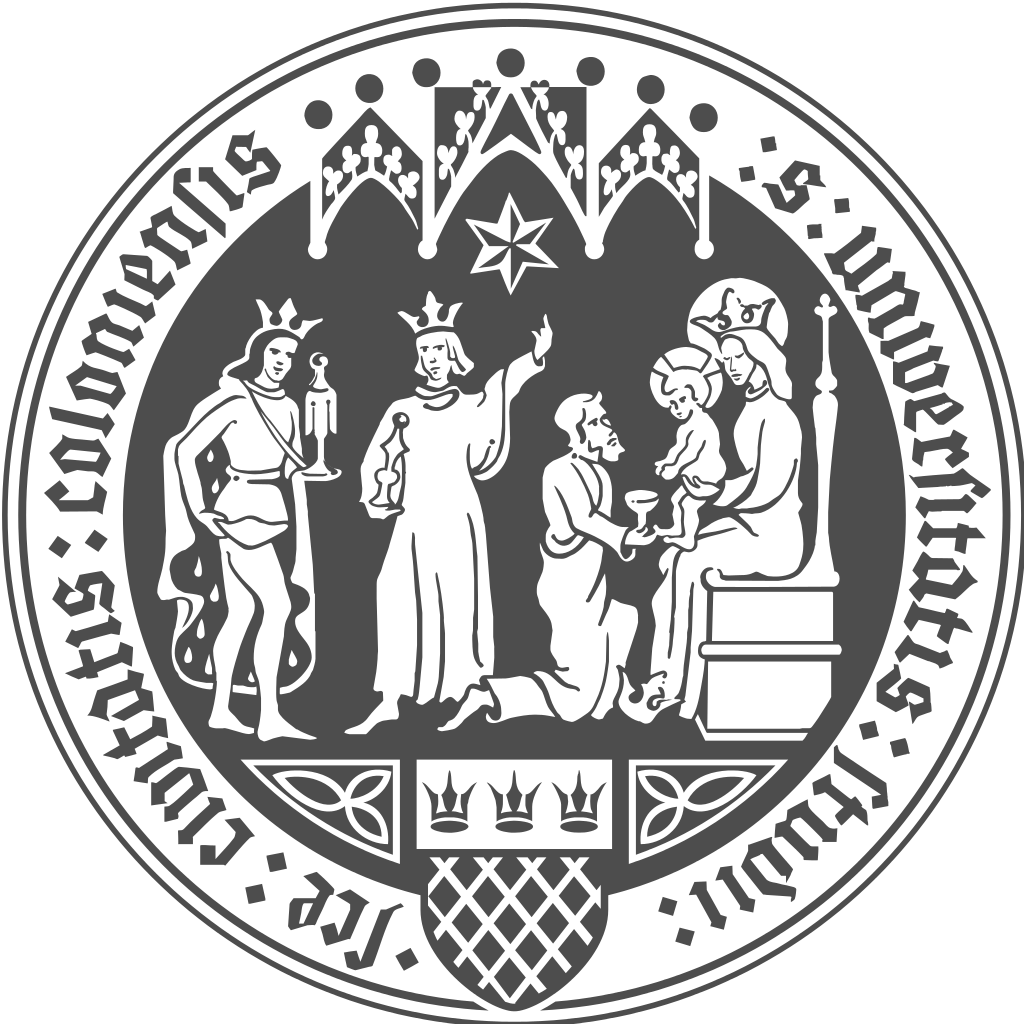
(c) Brief data preparation details (how your data were created from the raw data) and key charts. Details can be provided in an Appendix.

(d) Data analytics: Analytical methods applied (with sufficient detail and screenshots; use Appendix if needed) and appropriate performance evaluation (proper choice of measures, benchmarking).

(e) Conclusions (advantages and limitations) and business recommendations

Wordy Title - Even wordier subtitle

Bachelor Thesis / Master Thesis / Seminar Paper



Authors: Lukas Tempfli (Student ID: 7367097)

Moritz Danhausen (Student ID: 7369413)

Niklas Nesseler (Student ID: 7367375)

Robin Kirch (Student ID: 7364580)

Sven Dornbrach (Student ID: 7364484)

Team Name: ………………..

Supervisor: Univ.-Prof. Dr. Wolfgang Ketter

Co-Supervisor: Karsten Schroer

Department of Information Systems for Sustainable Society

Faculty of Management, Economics and Social Sciences

University of Cologne

July 21, 2021

# Eidesstattliche Versicherung

Hiermit versichere wir an Eides statt, dass wir die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Lukas Tempfli, Moritz Danhausen, Niklas Nesseler, Robin Kirch, Sven Dornbrach

Köln, den 21.07.2021

INHALTSVERZEICHNIS

|  |  |
| --- | --- |
| SUMMARY  TASK 2: One-page executive summary: summarizes the entire report for a non-technical manager (the business problem, data, the analytics solution, implications and recommendations)  This research project is based on real world data of the year 2017 from "Blue Bikes", which is a bike sharing firm based in Boston and focuses on the prediction of total bike usage in Boston for the next hour. In the following we will be acting as official employees of an outsourcing firm, which was commissioned to analyse the bike demand in Boston. The result are reported to a manager of Blue Bikes  To precisly predict the total bike usage marks a high value process, because it can minimize the cost of assest and increase the total value of the workflow. This is eminent, because to many bikes result in a not valueable utilatization over all bikes, which implifies that we could have save money regarding investment in the beginning. However the oppisit case is even worse, because this means we are lacking bikes and are not fully utilizing the market potential and could lead in bad user reputation, if they want to get a bike but they are not avaible.  The data we focus on is from 2017 and includes following attributes: …. This means we know exactly which trip in regards of duration, start & stop location, wheater, temperature were made.  Our analyse results in….  Implications we faced were…  We further would recommend…  Our target is to predict the demand of total bike usage for the next hour.  Our research is lies on data of the year 2017 and we will make a comprehensive analysis, in which we use typical data analysis and machine learning approaches to make use of the data to monitize and optimize the operations of "Blue Bikes". The overall topic is smart mobility systems and how we can use data in impactful ways to address pressing societal issues. In terms of bike sharing the addressed societal issue are the reduce of greenhouse gas emissions, reduce of pollution as health risk for urban population, reduce of (fatal) road accidents and to create a more efficient road transport infrastructur. We will cover the following task:   * **Task 1) Data Collection & Preparation**:   + Cleaning of datasets for use in later analysis stages * **Task 2) Descriptive Analysis**:   + Demonstrate temporal demand patterns and seasonality   + Demonstrate geographical demand patterns   + Define Key Performance Indicators which provides overview of current fleet operations * **Task 3) Predictive Analysis**:   + Forecast total system-level demand in the next hour   Popularity of station mit Heatmap  **3. Detailed report:**  **(a) Problem description (business goal and data science goal)**  In the business context of smart mobility services it is relevant to monitor the usage of the transport goods, in our case bicycle, to check whether there are enough bicycles available at each time of the day as well as for each station, so the customer is satisfied and relies more on a smart mobility solution. This is of high interest as good functioning smart mobility service reduces typical mobility issue that a big city has. These are traffic in city and place as well as pollution.  **(b) Data description**  The data on which the project relies is constructed of two different files.  Firstly, we were provided with operational raw data from “Blue Bikes” ("boston\_2017.csv") with the following attributes: start\_time (datetime), end\_time (datetime), start\_station\_id (int), end\_station\_id (int), start\_station\_name (str), end\_station\_name (str), bike\_id (int), user\_type (str).  Secondly, we were provided with hourly weather data for Boston ("weather\_hourly\_boston.csv"). The weather data ranges from the first January of 2015 to the second of January of 2020. It contains the following attributes: date\_time (“datetime”), max\_temp (float), min\_temp (float), precip (float).  For our task we are focusing mainly on these data sets. However we also use further information which was provided by the “Blue Bike” Website (378 docks) -> Stationen manuell löschen die noch nicht 2017 da sind, which is a data set which contain geographic information about the docking station. This will the used to for visualisation to get a feel where in Boston we have docking stations. The data sets contains following attributes: Number,Name,Latitude,Longitude,District,Public,Total docks.  **(c) Brief data preparation details (how your data were created from the raw data) and key charts. Details can be provided in an Appendix.**  **Was ist mit Duplikaten???**  Our data preparation for all three data sets is focused on two things. Firstly check for null values and secondly drop features we don’t need for our analysis.  **(d) Data analytics: Analytical methods applied (with sufficient detail and screenshots; use Appendix if needed) and appropriate performance evaluation (proper choice of measures, benchmarking).**  **(e) Conclusions (advantages and limitations) and business recommendations** | |
| John Doe *CONTENTS* | |
| Contents |  |
| 1 Section One | 1 |
| 1.1 Exemplary Table . . . . . . . . . . . . . . . . . . . . . . . . . . . | 1 |
| 1.2 Exemplary Section and Table Referencing . . . . . . . . . . . . . | 1 |
| 2 Section Two | 2 |
| 2.1 Exemplary Figure . . . . . . . . . . . . . . . . . . . . . . . . . . . | 2 |
| 2.2 Exemplary Figure Referencing . . . . . . . . . . . . . . . . . . . . | 2 |
| 3 Section X | 3 |
| 3.1 Exemplary Citation . . . . . . . . . . . . . . . . . . . . . . . . . . | 3 |
| A Appendix | 4 |
| References | 5 |

I

*LIST OF FIGURES* John Doe

# List of Figures

1 Exemplary Figure . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

II

John Doe *LIST OF TABLES*

# List of Tables

1. Exemplary Table . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

III

*1 SECTION ONE* 1 Section One

...

* 1. Exemplary Table

...

Table 1: Exemplary Table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Id | Col 1 | Col 2 | Col 3 | Col 4 | Col 5 |
| 1 | Col 1 | Col 2 | Col 3 | Col 4 | Col 5 |
| 2 | Col 1 | Col 2 | Col 3 | Col 4 | Col 5 |
| 3 | Col 1 | Col 2 | Col 3 | Col 4 | Col 5 |
| 4 | Col 1 | Col 2 | Col 3 | Col 4 | Col 5 |
| 5 | Col 1 | Col 2 | Col 3 | Col 4 | Col 5 |

* 1. Exemplary Section and Table Referencing

See Table 1 in Section 1.1 for details.

1. *SECTION TWO* John Doe

# Section Two

...

2.1 Exemplary Figure

...

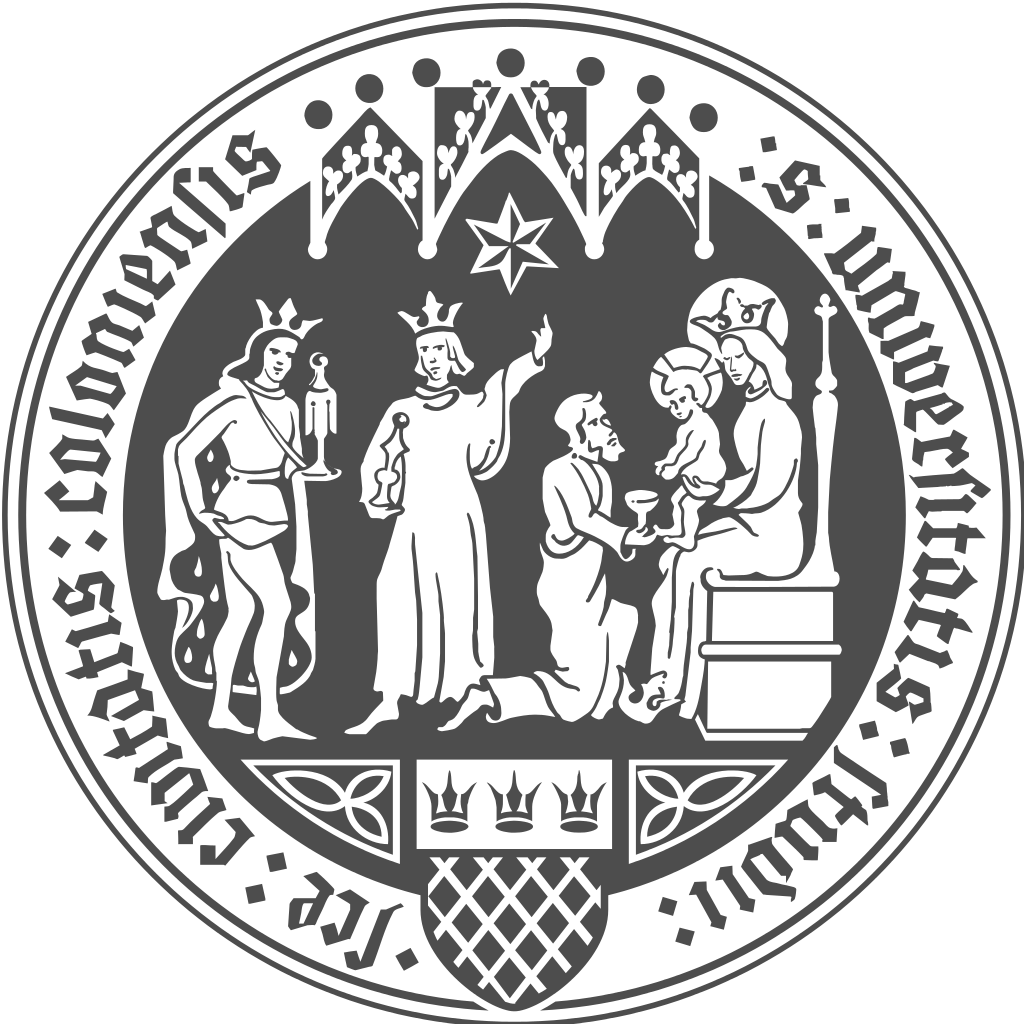


Figure 1: Exemplary Figure

2.2 Exemplary Figure Referencing

See Figure 1 for details. Additional information can be found in the footnote [[1]](#footnote-1).

2

*3 SECTION X*

# Section X

3.1 Exemplary Citation

In this research we follow Ketter et al. (2016)...

PowerTAC is an example of a multiagent competitive gaming platform (Ketter et al., 2016).

*A APPENDIX* John Doe

# A Appendix

...

4

*REFERENCES*

# References

Ketter, W., Peters, M., & Collins, J. (2016). A Multiagent Competitive Gaming Platform to Address Societal Challenges. *MIS Quarterly*, *40*(2), 447–460. doi:

10.25300/MISQ/2016/40.2.09

1. Image taken from [https://en.wikipedia.org/wiki/File:Siegel\_Uni-Koeln\_(Grau) .svg.](https://en.wikipedia.org/wiki/File:Siegel_Uni-Koeln_(Grau).svg) [↑](#footnote-ref-1)