Estimate of resource needs of the Final Paper

for the CEU MSc in Business Analytics program

Gábor Horváth

2017

CEU | CENTRAL
EUROPEAN
UNIVERSITY

1 Human resource needs

1.1 Stakeholders & people to involve

As this final paper is a pet project, there is no actual business management behind the requirements, therefore no business stakeholders are identified and involved. The completion of the project requires feedback and guidance from the mentor (Gergely Daroczi), but no other person (or role) needs to be involved.

1.2 Dedication for the project

There are no additional time dedication requirements identified above the requirements stated by CEU in the Final Project description document.

1.3 Training requirements

As stated earlier no additional organized / official training requirements are required above the trainings received during the courses in the program. There are tools and techniques used to fulfill the project, which have not been described in the program at CEU. There are several useful user manuals available on the webpages of the creators of the tools, which would enable the use of these tools and resources for any student who have been part of the program.

2 Software and data resource needs

2.1 Tools & resources used

As stated earlier, fulfilling the completion need for the project the following tools are planned to be used:

- Programming language:
  - R: https://www.r-project.org/
- IDE for the programming language:
  - RStudio: https://www.rstudio.com/
- Documentation is created using:
  - knitr: https://yihui.name/knitr/
  - MiKTeX: https://miktex.org/
  - ReporteRs: https://cran.r-project.org/web/packages/ReporteRs/index.html
- Data visualization:
  - ggplot2: http://ggplot2.org/
- Data manipulation:
  - access2csv: https://github.com/AccelerationNet/access2csv
  - dtplyr: https://cran.r-project.org/web/packages/dtplyr/index.html
- Project plan / task management:
  - Buckets: https://www.buckets.co/
- Source code repository:
  - GitHub: https://github.com/

Note: The list above do not contain the list of all the tools and packages used to create the project, but the full list will be provided in the source code.

2.2 Data sources

As stated earlier, the project will use the following data provided by multiple US government agencies:

- Federal Aviation Administration: Wildlife Strike Database
- United States Department of Transportation: Bureau of Transportation Statistics

Note: In case data enrichment would be required for the successful risk modelling, additional data sources might be used as well. These possible additional data sources will be listed in the Final Paper.

3 Contributors

Student: Gábor Horváth
Mentor: Gergely Daróczi