**Project** **Proposal**

**Project Title:** Filtering, Analysis, prediction, and final data visualization of Passenger flights (excluding small private aircraft with no details) data during the COVID-19 pandemic

**Student Name:** Xingyu Wang

**Student ID:** 2198771

**Supervisor Name:** Dr. Paul Blain Levy

**Project Aim**

The aim of this project is to use Crowdsourced air traffic data from the OpenSky Network, analyze it and experiment with selection algorithms to get the most accurate predicted departure and arrival times, and ultimately visualize the data with a sophisticated third-party visualization package in Python.

The reasons for choosing this project are large volumes of data from commercial aviation are collected and stored in databases daily to provide a better understanding of the entire flight life system. The entire flight ecosystem Efforts are being stepped up to extract useful knowledge from flight data generated by sensors and the Internet of Things In this regard, data science is a key tool for understanding the domain, managing data, and applying it to the environment.

**Project Objectives**

Objectives1: Filtering the database to obtain flight data for regional airliners and beyond.

Objectives2: Create algorithmic prediction models for experimentation to get the most accurate prediction data.

Objectives3: Visualization of experimental results and filtered data using the Matplotlib library in the Python library.

**Knowledge to be learned**

* Python
* Data analysis
* Data visualization
* Algorithm modelling

**Project Plan**

表格

描述已自动生成

**Risk and contingency plan**

The biggest challenge of this experiment is the need to obtain the optimal algorithm through a large number of experiments. If too much time is required to make the next step of the plan impossible, the prediction time will be obtained by referring to data from others' papers to directly select an algorithm. If the experimental data is ultimately unavailable, go to Plan B: Create a web page to implement historical flight queries and data visualization for flights.

**Hardware/Software Resources**

The majority of this project will be completed with an I7-9750H processor and a GTX1650 graphics card configuration. Any data that cannot be processed may need to be submitted to the school lab for processing. The code will be written using Visual Studio Code.

**Data**

Martin Strohmeier, Xavier Olive, Jannis Lübbe, Matthias Schäfer, and Vincent Lenders “Crowdsourced air traffic data from the OpenSky Network 2019–2020"*Earth System Science Data* 13(2), 2021 <https://doi.org/10.5194/essd-13-357-2021>

**References**

[1] Carvalho, Leonardo, et al. "On the relevance of data science for flight delay research: a systematic review." *Transport Reviews* 41.4 (2021): 499-528.