

# Configuration Manual

MSc Research Project  
Data Analytics

Nicholas Stancill  
Student ID: X22225668

School of Computing  
National College of Ireland

Supervisor: Dr Vladimir Milosavljevic

**National College of Ireland**  
**MSc Project Submission Sheet**  
**School of Computing**



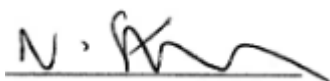
**Student Name:** Nicholas Stancill  
**Student ID:** X22225668  
**Programme:** Master of Science in Data Analytics      **Year:** 2025  
**Module:** MSc Research Project  
**Lecturer:** Dr Vladimir Milosavljevic  
**Submission Due Date:** 15/09/2025  
**Project Title:** Predicting Competitive Dynamics in Formula 1: Modelling the Impact of FIA Aerodynamic Regulations Using Machine Learning and Simulation

**Word Count:** 338    **Page Count:** 5

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:**



**Date:** 14/09/2025

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
<b>Attach a Moodle submission receipt of the online project submission,</b> to each project (including multiple copies).	<input type="checkbox"/>
<b>You must ensure that you retain a HARD COPY of the project,</b> both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

<b>Office Use Only</b>	
Signature:	
Date:	

Penalty Applied (if applicable):	
----------------------------------	--

# Configuration Manual

Nicholas Stancill  
Student ID: X22225668

## 1 Experimental Environment

This section outlines the development environment used to conduct the experimental work for this project.

- **Platform:** Anaconda Navigator 2.5.0
- **Operating System:** Windows 11 Home (64-bit)
- **Python Version:** 3.11.4
- **Jupyter Notebook Version:** 6.5.4
- **Execution Environment:** Jupyter Notebook launched via Anaconda Navigator
- **Environment Used:** Base (no virtual environments or containers)

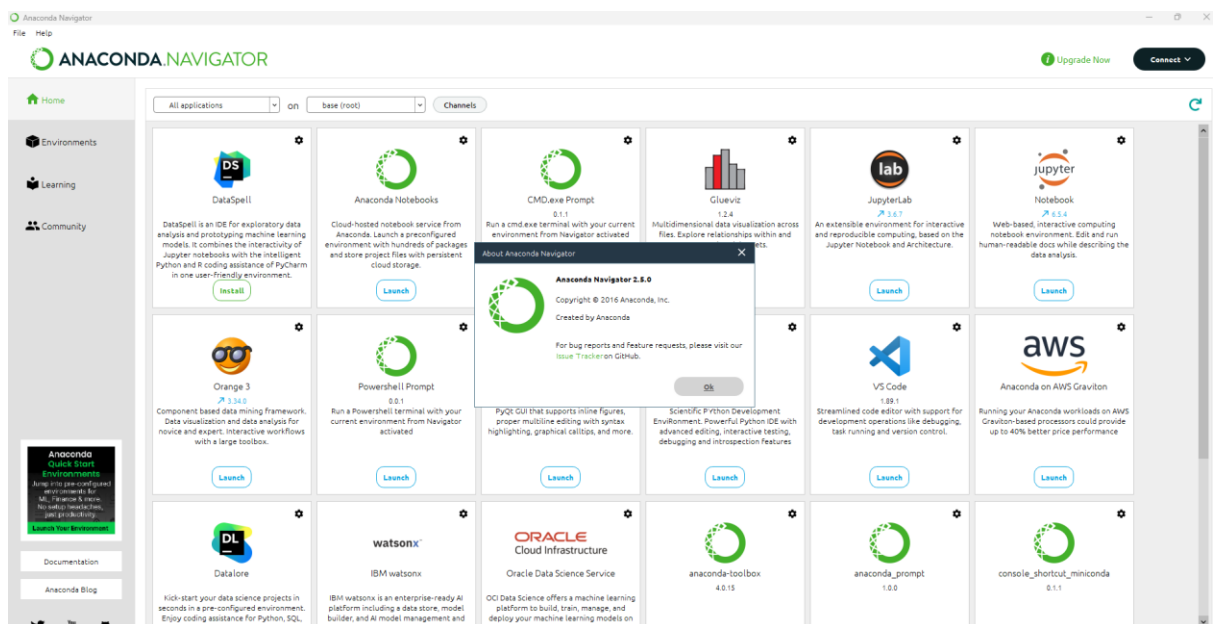



Fig 1 - Anaconda Navigator showing version 2.5.0 and Jupyter Notebook Version 6.5.4

System type 64-bit operating system, x64-based processor

Pen and touch No pen or touch input is available for this display

**Related links** [Domain or workgroup](#) [System protection](#) [Advanced system settings](#)

 Windows specifications Copy ^

Edition Windows 11 Home

Version 23H2

Installed on 17/08/2023

OS build 22631.4751

Experience Windows Feature Experience Pack 1000.22700.1055.0

[Microsoft Services Agreement](#)

[Microsoft Software Licence Terms](#)

**Fig 2 – Operating system**

```
In [1]: 1 import sys
        2 print(sys.version)

3.11.4 | packaged by Anaconda, Inc. | (main, Jul 5 2023, 13:38:37) [MSC v.1916
64 bit (AMD64)]
```

**Fig 3 – Python 3.11.4 output in notebook**

The environment was used without any external editors or cloud services. Execution was performed locally.

## 2 Required Libraries

Library	Version	Purpose
pandas	1.5.3	Data manipulation and processing
numpy	1.24.3	Numerical operations
matplotlib	3.8.3	Static visualisations
seaborn	0.12.2	Statistical data visualisation
plotly	5.18.0	Interactive plots
scikit-learn	1.1.3	Machine learning and data preprocessing
scipy	1.10.0	Scientific/statistical computations
os	Built-in	File and directory operations

All libraries were available in the Anaconda base environment, with a few installed using either pip or conda.

```
In [2]: 1 import importlib, pandas as pd, numpy as np, matplotlib, seaborn, plotly, scipy, sklearn
2 print("pandas", pd.__version__)
3 print("numpy", np.__version__)
4 print("matplotlib", matplotlib.__version__)
5 print("seaborn", seaborn.__version__)
6 print("plotly", plotly.__version__)
7 print("scipy", scipy.__version__)
8 print("scikit-learn", sklearn.__version__)
9

pandas 1.5.3
numpy 1.24.3
matplotlib 3.8.3
seaborn 0.12.2
plotly 5.18.0
scipy 1.10.0
scikit-learn 1.1.3
```

Fig 4 – Installed Python libraries and versions

### 3 Dataset and Execution

The dataset used for the project was:

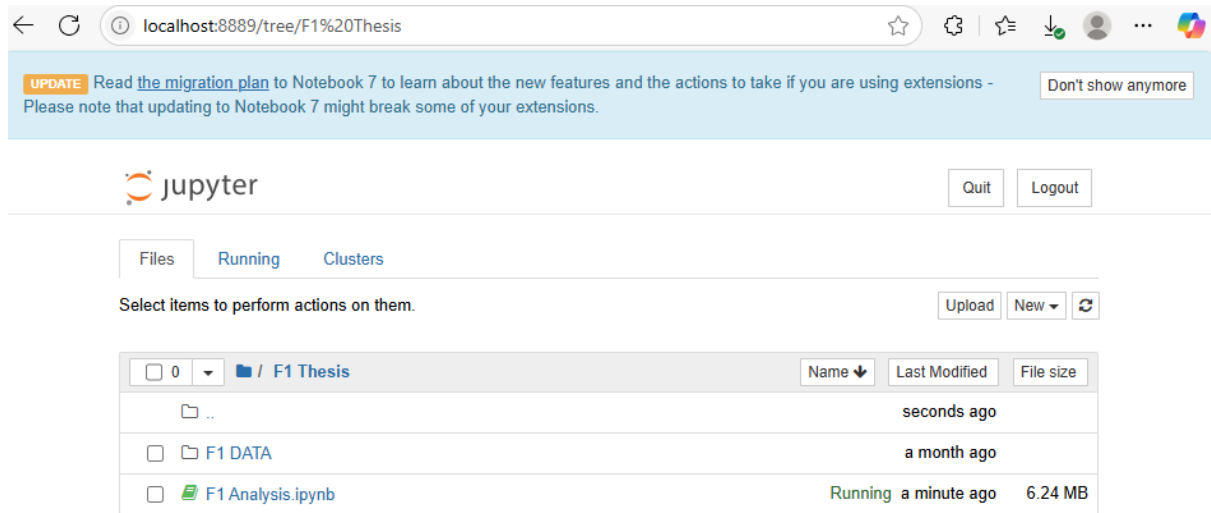
- **Title:** Formula 1 World Championship (1950 - 2020)
- **Author:** Rohan Rao
- **Source:** <https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020>

#### Usage Notes:

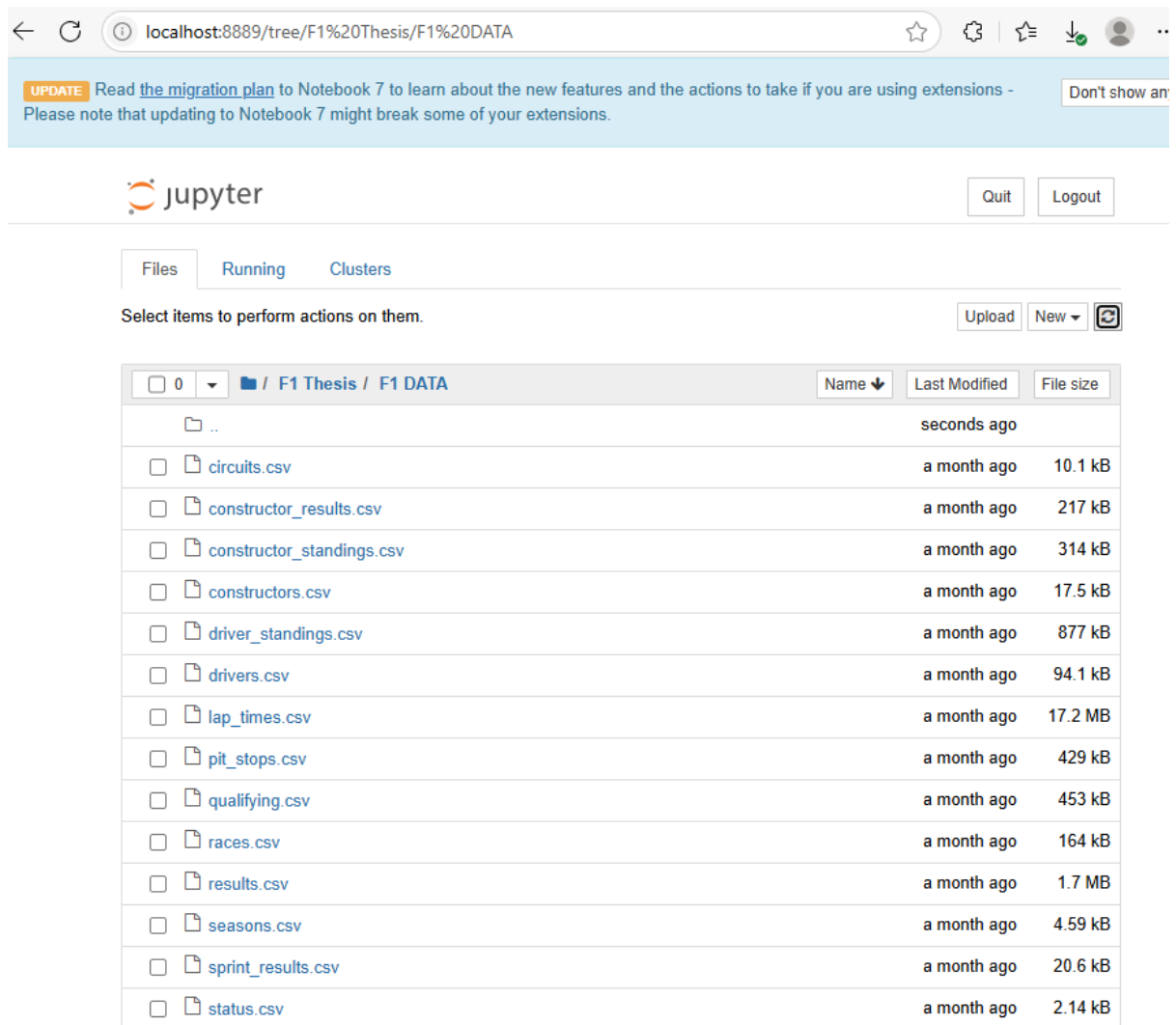
- The dataset folder has been renamed to F1 DATA and placed inside the project directory F1 Thesis.
- The notebook file used for this project is named F1 Analysis.ipynb and resides in the same folder.
- The notebook was executed sequentially from top to bottom using the "Run All" option in Jupyter Notebook.
- Outputs generated include plots, tables, and machine learning results, all displayed inline within the notebook.
- The dataset files (e.g. races.csv, drivers.csv, etc.) were referenced using relative paths such as: `base_path = "F1 DATA"`

#### Reproducibility Notes:

- It is important to use the exact versions of the libraries listed to avoid incompatibilities.
- Ensure that all files are in the same working directory.
- Confirm the Python kernel is set to 3.11 (base) before execution.



**Figure 5 – Project folder structure inside Jupyter**



**Figure 6 – Dataset folder contents (CSV files)**

The screenshot shows a Jupyter Notebook titled "F1 Analysis (autosaved)" running on a Python 3 (ipykernel) environment. The code in the notebook is as follows:

```
5 import pandas as pd
6 import numpy as np
7
8 # Visualisation Libraries
9 import matplotlib.pyplot as plt
10 import seaborn as sns
11 import plotly.express as px
12
13 # Statistical testing Libraries
14 from scipy.stats import ttest_ind, f_oneway, t
15
16 # Machine Learning Libraries
17 from sklearn.model_selection import train_test_split, KFold, cross_val_score
18 from sklearn.ensemble import RandomForestRegressor
19 from sklearn.preprocessing import LabelEncoder
20
21 # Model evaluation metrics
22 from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
23
24 # Base path for data files
25 base_path = "F1 DATA"
26
```

### 3.1 Initial Check of Dataset

In [2]:

```
1 # List all CSV files available in the specified data directory
2 available_files = [
3     file for file in os.listdir(base_path)
4     if file.endswith(".csv")
5 ]
6
7 # Display the list of available CSV files
8 available_files
9
```

Out[2]:

```
['circuits.csv',
'constructors.csv',
'constructor_results.csv',
'constructor_standings.csv',
'drivers.csv',
```

**Fig 7 – Jupyter Notebook showing imported libraries, dataset path definition, and file discovery from the F1 DATA folder**

## References

Kaggle (2024) *Formula 1 World Championship (1950–2024)* [Dataset]. Available at: <https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020>