**MSc in Data Analytics (SB+) - Sept 2023 - 2023 - YR1**

Author: Jose Maria Rico Leal

e-mail: sba23021@student.cct.ie

Student ID: sba23021

GitHub Link:

Irish transport sector

## **Abstract**

*Empty*

[**Abstract** 1](#_Toc151929784)

[**1.** **Introduction** 3](#_Toc151929785)

[**2.** **Programming** 4](#_Toc151929786)

[2.1. Programming 4](#_Toc151929787)

[2.2. Data structures 4](#_Toc151929788)

[2.3. Documentation 4](#_Toc151929789)

[2.4. Testing and optimisation 4](#_Toc151929790)

[2.5. Data manipulation 4](#_Toc151929791)

[**3.** **Statistical Analysis** 4](#_Toc151929792)

[3.2. Confidence interval 4](#_Toc151929793)

[3.3. Inferential statistics 4](#_Toc151929794)

[*3.3.1.* *Parametric* 4](#_Toc151929795)

[*3.3.1.1.* *T-test* 4](#_Toc151929796)

[*3.3.1.2.* *Anova* 4](#_Toc151929797)

[*3.3.1.3.* *Chi-squared test* 4](#_Toc151929798)

[*3.3.2.* *Non-parametric* 4](#_Toc151929799)

[*3.3.2.1.* *Kruskal-Wallis* 4](#_Toc151929800)

[*3.3.2.2.* *U-Mann Whitman* 4](#_Toc151929801)

[3.4. Outcome and challenges faced 4](#_Toc151929802)

[**4.** **ML** 5](#_Toc151929803)

[4.1. ML supervised learning 5](#_Toc151929804)

[4.2. Sentiment analysis 5](#_Toc151929805)

[4.3. Comparing Supervised, Unsupervised and semi-supervised ML models. 5](#_Toc151929806)

[4.4. Table and conclusions 5](#_Toc151929807)

[**5.** **Data Preparation and Visualisation** 5](#_Toc151929808)

[5.1. Data acquistion 5](#_Toc151929809)

[5.2. EDA methodology 5](#_Toc151929810)

[5.3. Visualisations 5](#_Toc151929811)

[5.4. Dashboard 5](#_Toc151929812)

[**6.** **Conclusion** 5](#_Toc151929813)

[**7.** **References** 5](#_Toc151929814)

# **Introduction**

# **Programming**

# Programming

# Data structures

# Documentation

# Testing and optimisation

# Data manipulation

# **Statistical Analysis**

* 1. Descriptive statistics

# Confidence interval

# Inferential statistics

# *Parametric*

# *T-test*

# *Anova*

# *Chi-squared test*

# *Non-parametric*

# *Kruskal-Wallis*

# *U-Mann Whitman*

# Outcome and challenges faced

# **ML**

# ML supervised learning

# Sentiment analysis

# Comparing Supervised, Unsupervised and semi-supervised ML models.

# Table and conclusions

# **Data Preparation and Visualisation**

# Data acquistion

# EDA methodology

# Visualisations

# Dashboard

# **Conclusion**

# **References**