**Data Pipeline  
An Apache Airflow DAG and Python Script**

**1. Project Description:**This is a Data pipeline that automates the downloading, cleaning, and uploading of data in two different data bases depending on the version: MS SQL Server the Python script, and MySQL in the Apache Airflow Version. The purpose of this project was the gain hands-on experience in creating data pipelines and utilising the Apache Airflow tool.

**2. Problem Statement:**This project aimed to solve the lack of hands-on experience in creating data pipelines, and lack of comprehension of Apache Airflow and its ability. Having finished my data cleaning and visualisation project “Aim\_Lab”, and frequently seeing ‘Apache Airflow’ and ‘data pipelines’ in job descriptions, it was the natural next step in my learning path.   
  
**3. Solution Implementation:**

**4. Results and Impact:**  
 **5. Technical Skills Demonstrated:  
  
  
6. GitHub Repository:**

**7. Key Takeaways:**

This Project is a data pipeline. With one click of a button, the data will download, clean and format, and upload into different databases depending on the version. MS SQL Server in the python version, and MYSQL in the Apache Airflow version.   
  
The main objective was to gain experience in creating data pipelines, to utilise and learn the tools, and understand the potential of data pipelines.   
  
This project has two versions, one in Python and the other in Apache Airflow. I wanted to learn how to make data pipelines in Apache Airflow, as it is a scalable and valuable tool. And then utilised the opportunity to push the data into two different databases.