National College of Ireland

Configuration Manual

MSc Research Project
Cloud Computing

Muhammad Osama Khan
Student ID: x21242887

School of Computing
National College of Ireland

Supervisor: Diego Lugones

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Muhammad Osama Khan
Student ID:	x21242887
Programme:	Cloud Computing
Year:	2023
Module:	MSc Research Project
Supervisor:	Diego Lugones
Submission Due Date:	14/08/2023
Project Title:	Configuration Manual
Word Count:	643
Page Count:	4

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:	Muhammad Osama Khan
Date:	12th August 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

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

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

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Muhammad Osama Khan
x21242887

Google Colab is a free cloud based platform which uses web browser to write and execute Python Code. The environment is similar to to Jupyter Notebooks. This manual will provide a guidance on how to run provided python code in Google Colab.

Prerequisites

1) A Google account.
2) A web browser (e.g., Google Chrome, Mozilla Firefox).

1 Access Google Colab

1) Open your web browser.
2) Navigate to Google Colab website.
3) If prompt will appear then Sign in using your Google account credentials.

2 Create a New Notebook

1) On Colab homepage you can click on the + NEW NOTEBOOK button.
2) New NOTEBOOK will be created.
3) Upload provided notebook(x21242887-Research-Code) which is inside "x21242887-ICT Solution Artefact" zip folder that is submitted on MyMoodle.

3 Familiarize Yourself with the Interface

The Colab interface consists of:

Menu bar: It contains various options like File, Edit, View, Insert, etc.
Toolbar: It provides features for commonly used actions such as saving, adding code or text cells, etc.
Main Workspace: This is where the code and text cells will display.

4 Uplaod Event CSV file

1) There are few icons on the left-hand sidebar of Colab inerface. Click on the folder icon, which represents the file tab.
2) Click on the upload button after opening the file tab. I have placed "events.csv" file

in "x21242887-ICT Solution Artefact" zip folder that is submitted on MyMoodle. Select that "event.csv" file and uplaod it.
3) CSV file will be uploaded to the main directory of your Colab Environment after.
4) Progress circle will be visible in front of the file, once complete then it will be listed in Files tab.

5 Section 5

1) Click on the Play Button of first cell and wait until it will turn green.
2) When the cell will turn green then you should move to the next cell and press play button again until last cell.

6 Code Flow

Code is divided into multiple cells, one cell is executing only one step at a time. Below are the details of the cells.

6.1 Import Libraries

Press play button and Libraries will be imported to the Notebook, Figure 1.

```
!pip install pandas faker
!pip install openpyxl
!pip install matplotlib
```

Figure 1: Import Libraries

6.2 Transform Event File

Event.csv file will be transformed in to output.csv by executing the next code cell. Output.csv will be placed in the environment directory, Figure 2.

6.3 Generate Client Application Data

The client application data will be generated Using Output.csv file. The result will be saved in client_billing_data.xlsx file, which will then placed in Environment directory, Figure 3.

6.4 Plot the result on Barchart

Date generated from previous code cell will be plotted to Barchart, which will show Cost per Month.

```
[ ]  import pandas as pd
     from dateutil.relativedelta import relativedelta

     # Load the data
     df = pd.read_csv('events.csv')

     # Convert the 'event_time' column to a datetime
     df['event_time'] = pd.to_datetime(df['event_time'])

     # Set 'event_time' as the index
     df.set_index('event_time', inplace=True)

     # Repeat the data four times to cover 2 years
     df_list = []
     for i in range(4):  # repeat four times for 2 years
         df_temp = df.copy()
         df_temp.index = df_temp.index + pd.DateOffset(months=5*i)
         df_list.append(df_temp)

     df = pd.concat(df_list)

     # Resample by hour and count events
     df_resampled = df.resample('H').size()
     df_resampled *= 10
     # Create output dataframe
     df_output = pd.DataFrame(df_resampled, columns=['number_of_actions'])

     # Save to a new csv file
     df_output.to_csv('output.csv')
```

Figure 2: Transform Event File

6.5 Predict 4 years Forecast Data

This code cell will predict 4 years Forecast Billing and resource utilization Data for Client Application and save the results in client_billing_data_forecast.xlsx file.

6.6 Plot the result on Barchart

This code section will plot billing forecast results generated from the previous section to a Barchart.

6.7 Optimize and Predict 4 years of Forecast Data

This code cell will predict 4 years Forecast Billing and resource utilization Data after performing optimization for Client Application and save the results in client_billing_data_forecast_cheap_resources.xlsx file.

6.8 Plot the result on Barchart

This code section will plot optimized billing forecast results generated from the previous section to a Barchart.

6.9 Plot the result of Simple Forecast and Optimized Forecast on Line Chart

This code section will plot the results of Simple Forecast and Optimized Forecast on Line Chart for the comparison.

```
[ ] import random
    from faker import Faker
    import pandas as pd

    # Load the CSV data
    df_actions = pd.read_csv('output.csv')
    df_actions['event_time'] = pd.to_datetime(df_actions['event_time'])
    df_actions.set_index('event_time', inplace=True)


    # Create a Faker instance
    fake = Faker()

    # Define AWS services, types, and costs
    services = {
        'AzureCompute': {
            'types': ['B1 Series', 'B1ms Series', 'B2s Series'],
            'costs': {'B1 Series': 0.016, 'B1ms Series': 0.028, 'B2s Series': 0.056}
        },
        'RDS': {
            'types': ['db.t2.micro', 'db.t2.small', 'db.t2.medium'],
            'costs': {'db.t2.micro': 0.018, 'db.t2.small': 0.036, 'db.t2.medium': 0.073}
        },
        'S3': {
            'types': ['GET', 'PUT', 'COPY', 'POST', 'LIST'],
            'costs': {'GET': 0.0004, 'PUT': 0.0005, 'COPY': 0.0005, 'POST': 0.0005, 'LIST': 0.005}
```

Figure 3: Generate Client Application Data

All the files which will be generated by the result of execution of each cell can be downloaded form the main directory of Colab environment. Furthermore, analysis can be done using the downloaded files.