

**Data and Web Analytics**

COMP1810

Full Name: DINH NGUYEN HOANG ANH

Class: COS1203

Student ID: 001405962 – GCS230495

Submission date:

Deadline Date: 7/8/2025

Table of Contents

[I. Introduction 2](#_Toc202177263)

[II. Task 1 2](#_Toc202177264)

[2.1 Task a 2](#_Toc202177265)

[2.2 Task b 2](#_Toc202177266)

[2.3 Task c 2](#_Toc202177267)

[III. Task 2 2](#_Toc202177268)

[3.1 Task d 2](#_Toc202177269)

[3.2 Task e 2](#_Toc202177270)

[IV. References 2](#_Toc202177271)

[V. Appendix 2](#_Toc202177272)

# Introduction

Data analysis is an important subject that teaches us how to turn basic numbers and facts into useful stories and ideas. This work includes steps such as gathering, cleaning, and looking at data, and then using tools to solve problems. Because this skill can be used in so many ways, it is now very popular and needed in many fields. For example, it is used in business and banking to guess what will happen in the future and manage risk. It's also used in marketing and retail to better understand customers.

# Task 1

It can be seen from the Web Analytic\_Dataset.csv that the dataset has 250 data points and 13 variables. Additionally, there are 3 category columns including Source / Medium, Bounce Rate and Conversion Rate (%). About numerical columns, there are Year, Month of the year, Users, New Users, Sessions, Pageviews, Transactions, Revenue and Quantity Sold (9 columns). Therefore, the Avg. Session Duration format in time-series.

About the issues, bounce rate and Conversion rate (%) contain number in each value. However, the Bounce rate and Conversion rate stored as character. As a result, we need to convert data types of these columns.

## Task a

Table

chart

## Task b

## Task c

# Task 2

## Task d

## Task e

# References

# Appendix

Coursework file: