IT3382 ADVANCED DATA VISUALISATION

Assignment (40%)

AY2024 Semester 2

Table of Contents

Introduction

By the end of this assignment, you will be able to:

- Detect and understand the stories within datasets and extract insights from the given dataset
- Effectively present data visually to enhance audience comprehension of findings and insights.
- Design and create dynamic and interactive data visualisations using scripting languages and Python libraries
- Act as a data-driven visual storyteller for optimal presentation of trends, patterns and insights.

Tasks Involved in Assignment

Below is the list of tasks to be performed in this assignment.

S/N	Task	Submission
1.	Develop a dynamic and interactive dashboard using D3, HTML, CSS and JavaScript (20 marks) Present the business insights using D3 Dashboard (5 marks)	A D3 dashboard that effectively presents data visually to enhance audience comprehension of findings and insights. A video file in mp4 format or a YouTube link to your presentation.
2.	Visualise Data Insights using Seaborn (15 marks)	Python Notebook file

In this **individual** assignment, you are required to develop interactive and effective data visualisations and integrate the visualisations into a dashboard using D3, HTML, CSS, JavaScript and Seaborn library. The datasets of this assignment are CSV files.

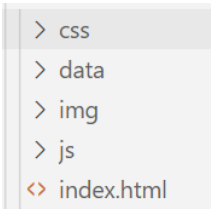
The base marks of this assignment are **40 marks,** and it constitutes **40%** of your total ICA marks for this competency unit.

Copy work from other people or the internet is strictly prohibited. If found, it will be considered a case of plagiarism and is subject to disciplinary actions

Submission Format and Mode

Below are the required deliverables for this assignment.

1. A D3 dashboard which contains **3-4** visualisations. Below is the folder structure you are required to follow to have better organisation of your source codes.



2. A recorded video of you presenting the dashboard (you may submit a video file in mp4 format or a YouTube link to your video. The duration of the video should not be more than 5 minutes.
3. Jupyter Notebook source file with file extension, .ipynb for Task 2.

Please refer to **Annex A** for detailed assessment rubrics of Task 1. Task 2 will be marked based on the correctness of the Python codes.

Late Submission for Assignment

Please be reminded to submit all the deliverables via Brightspace **by the end of Week 9, 15 December 2024 (Sunday) at 2359hrs.**

No. of Calendar Days	**Penalty**
If the submission is <= 5 working days	Cap at 50% (of the base marks)
If the submission is > 5 working days	0 marks will be awarded

Guidelines for Policy

- Do not use AI tools for plagiarism or cheating without proper citation or acknowledgement.
- Use AI tools ethically to supplement studies, clarify concepts, and assist with research.
- Cite sources and credit AI tools in academic work.
- Demonstrate understanding and application of AI-generated output when asked.

Task 1: Develop a dynamic and interactive dashboard on Commercial Property Sales Transactions using D3, HTML, CSS and JavaScript (25 marks)

You are tasked to visualise Commercial Building Prices using D3, HTML, CSS and JavaScript. The objective of this dashboard is to look at price distribution according to area as well as different commercial property types.

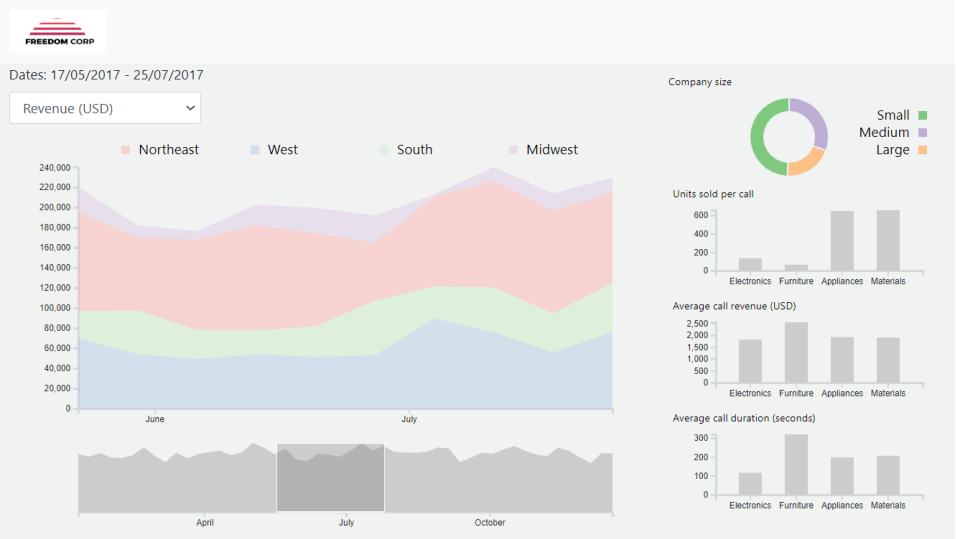
Data Source: **CommercialTrans_201910 to 202410.csv**

(Extracted from URA Commercial Property Transactions, Commercial Property Transactions - Property Market Information)

The dataset contains 3,556 rows of data about commercial property transactions with caveats lodged within the last 60 months. Caveats are legal documents lodged by purchasers with the Singapore Land Authority to register their legal interest in the property. Caveats are usually lodged by purchasers after the Option to Purchase is exercised or the Sales and Purchase agreement is signed.

Data Column
Project Name
Street Name
Property Type
Transacted Price ($)
Area (SQFT)
Unit Price ($ PSF)
Sale Date
Type of Area
Area (SQM)
Unit Price ($ PSM)
Tenure
Postal District
District Name
Floor Level

Your dashboard should contain 3-4 visualisations. In your dashboard, you should provide data filtering feature for your users. Please see below for an example of D3 dashboard extracted from Udemy course - Mastering data visualization in D3.js!



You are also required to record a video to present the business insights using the D3 dashboard you have created. Your presentation should be aided by the D3 features, no Microsoft PowerPoint slides are required.

Task 2: Visualise Data using Seaborn (15 marks)

To assess your competencies for the job role you have applied, the recruiting manager has given you a dataset and indicated the list of tasks to be completed in the Python Notebook template file.

Template File: assignment_task2_student.ipynb

Data Source: cardiac_disease.csv

This data set dates from 1988 and consists of four databases: Cleveland, Hungary, Switzerland, and Long Beach V. It contains 76 attributes, including the predicted attribute, but all published experiments refer to using a subset of 14 of them. The "target" field refers to the presence of heart disease in the patient. It is integer valued 0 = no disease and 1 = disease.

S/N	Variable	Explanation
1	age	Age in years
2	sex	1 = Male, 0 = Female
3	cp	Chest Pain Type - [0: asymptomatic, 1: atypical angina, 2: non-anginal pain, 3: typical angina]
4	trtbps	Resting Blood Pressure (mmHg)
5	chol	Cholesterol (mg/dL)
6	fps	Fasting Blood Sugar > 120 mg/dl - [1 = Yes, 0 = No]
7	restecg	Resting ECG - [0: showing probable or definite left ventricular hypertrophy by Estes' criteria, 1: normal, 2: having ST-T wave abnormality]
8	thal	Maximum Heart Rate Achieved (bpm)
9	exang	Exercise Induced Angina - [1 = Yes, 0 = No]
10	oldpeak	ST Depression Induced by Exercise
11	slope	Slope of the Peak Exercise ST Segment - [0: downsloping; 1: flat; 2: upsloping]
12	ca	Number of major vessels (0–3)
13	thal	[1 = normal, 2 = fixed defect, 3 = reversible defect]
14	target	[0 = disease, 1 = no disease]

Assignment Task 2: Visualise Data Insights using Seaborn (15 marks)

Sub Task 1: Data Preparation(2 marks)

Sub Task 2: Show Distribution of Different Numerical Variables (2 marks)

Sub Task 3: Show Distribution of Different Categorical Variables (2 marks)

Please refer to assignment_task2_student.ipynb for the complete list of tasks.

Annex A: Assessment Rubrics

Task 1: Develop a dynamic and interactive dashboard using D3, HTML, CSS and JavaScript					
Criteria	**Advanced**	**Proficient**	**Functional**	**Developing**	**Not Competent**
Data Representation **(5 marks)**	All the graphic variable types used are suited for the type and scale of the data they represent. ***(4.1 – 5 marks)***	Most of the graphic variable types used are suited for the type and scale of the data they represent. ***(3.1 – 4 marks)***	At least half of the graphic variable types used are suited for the type and scale of the data they represent. ***(2.1 – 3 marks)***	Only a few graphic variable types used are suited for the type and scale of the data they represent. ***(1.1 - 2 marks)***	No graphic variable types used are suited for the type and scale of the data they represent. ***(0-1 mark)***
Visualisation **(5 marks)**	All the visualisations address the required analysis concisely and clearly. ***(4.1 – 5 marks)***	Most of the visualisations address the required analysis concisely and clearly. ***(3.1 – 4 marks)***	At least half of the visualisations address the required analysis concisely and clearly. ***(2.1 – 3 marks)***	Only a few visualisations address the required analysis concisely and clearly. ***(1.1 - 2 marks)***	The visualisation did not address the required analysis concisely and clearly. ***(0-1 mark)***
Dashboard **(10 marks)**	The dashboard was well designed with excellent use of advanced tools in D3 features such as brushes, tooltips, and layout. ***(8.1 – 10 marks)***	The dashboard was designed with appropriate use of D3 features such as brushes, tooltips, and layout. ***(6.1 - 8 marks)***	The dashboard was designed with moderate use of D3 features such as brushes, tooltips, and layout. ***(4.1 - 6 marks)***	The dashboard was designed with minimal use of D3 features such as brushes, tooltips, and layout. ***(2.1 - 4 marks)***	The dashboard is not interactive and did not use any D3 features. ***(0 - 2 marks)***
Analysis **(5 marks)**	Concise analysis with comprehensive recommendations conveyed. ***(4.1 – 5 marks)***	Clear analysis with recommendations conveyed. ***(3.1 – 4 marks)***	Moderate analysis with recommendations conveyed. ***(2.1 – 3 marks)***	Brief analysis with irrelevant recommendations conveyed. ***(1.1 - 2 marks)***	Unclear analysis without recommendation. ***(0-1 mark)***