UDACITY Google spark

Palestine Launchpad Program

Data Analysis NanoDegree

Our agenda for today

1. Recap about the last session
2. Progress checks v. program timeline
3. Cohort 3 registration
4. Project 3 Rubric – Continue
5. Univariant analysis – Line Chart
6. Knowledge Share (NLP data processing)
7. Q/A

Palestine Launchpad
منصة فلسطين التقنية
Google

Recap

What did we talk about last session

UDACITY

Last Session we covered the following:

- Pareto Diagram: We into its significance and how to create one effectively.

- Project Three – Part 1: Exploratory Data Analysis Rubric: We discussed the expectations and requirements for the first part of your third project.

- Histogram in Seaborn: We explored how to create and interpret histograms using the Seaborn library in Python.

- Selenium Web Driver: We introduced Selenium for web scraping and browser automation.

Palestine Launchpad
منصة فلسطين التقنية
Google

Cohort 3 Registerstion

Regisieration is open

UDACITY

Palestine Launchpad
منصة فلسطين التقنية
Google.org

UDACITY Google spark

Cohort 3

We have **3 Specializations (Tracks)** and a total of **6 NDs**

Become a Data Scientist

1

Data Science
Programming for Data Science with Python

4

Data Science
Data Analyst

Become an AI Engineer

2

AI Engineering
AI Programming with Python and TensorFlow

5

AI Engineering
Intro to Machine Learning with TensorFlow

Become a Web Developer

3

Web Development
Intro to Programming

6

Web Development
Front End Web Developer

Cohort 3

We have **3 Specializations (Tracks)** and a total of **6 NDs**

Registration link Palestine Launchpad with Google - SPARK

You **must** pass this Nanodegree in order to be able to get another scholarship. Otherwise, you will be **blacklisted**

Remember, your acceptance into this scholarship program came at the expense of someone else who was not selected. If you are not fully committed, please do not deprive others of this valuable opportunity.

Palestine Launchpad
منصة فلسطين التقنية
Google

Progress checks v. program

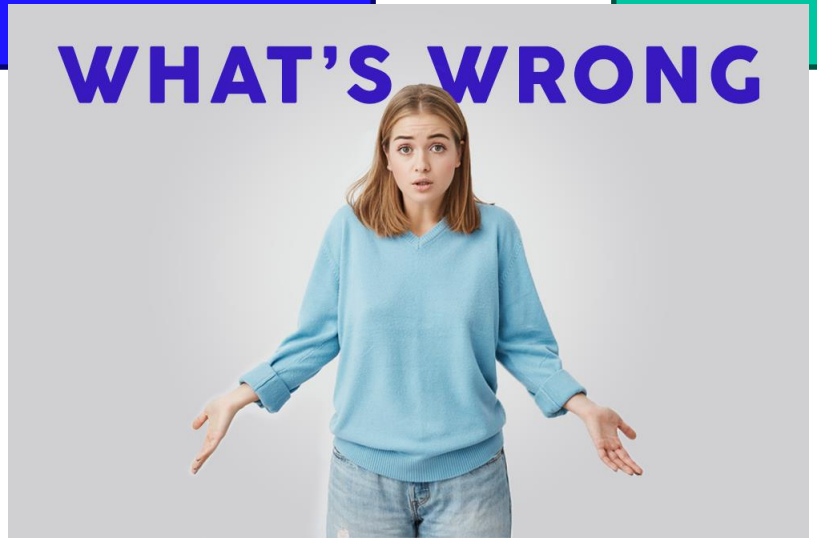
Constant is a key sucess

UDACITY

Off Track Studetns

Those who are not aligned with Weekly schedule

Students Who did not submit or pass The second project



Palestine Launchpad
منصة فلسطين التقنية
Google

Project Three : Communicate Data Findings

What is needed to make a successful submit

UDACITY

Project: Communicate Data Findings

Project Overview
This project has two parts that demonstrate the importance and value of data visualization techniques in the data analysis process.

- ✓ **In Part I, Exploratory data visualization**, you will use Python visualization libraries to systematically explore a selected dataset, starting from plots of single variables and building up to plots of multiple variables.

- ✓ **In Part II, Explanatory data visualization**, you will produce a short presentation that illustrates interesting properties, trends, and relationships that you discovered in your selected dataset. The primary method of conveying your findings will be through transforming your exploratory visualizations from the first part into polished, explanatory visualizations.

- ✓ **Project Due Date: Aug 27, 2024**

Choose Your Dataset

Below is a compiled list of the datasets you can choose from. However, you can explore other datasets that interest you:

- ❑ **Ford GoBike System Data(opens in a new tab)** (38 MB, CSV File)
- ❑ **Flights**
- ❑ **Loan Data from Prosper(opens in a new tab)** (82.5 MB, CSV File)
- ❑ **PISA Data - 2018**

Pareto Charts & 80-20 Rule - Clinical Excellence Commission (nsw.gov.au)

Univariate Exploration

- Histogram - required
- Bar Chart
- Count Plot

Bivariate Exploration

- Scatterplots - required
- Box Plots - required
- Clustered Bar Chart
- Heatmap

Multivariate Exploration

- Facet Plot - required
- Plot Matrix
- Scatterplot with multiple encodings

Project Three Rubric –PART ONE

Is the data explored systematically using a series of appropriate and varied visualizations?

Student has provided seven (7) exploratory data visualizations distributed over univariate, bivariate, and multivariate plots to explore many relationships in the data set.

For each exploration **category**, complete the required chart(s) and choose one additional visualization type.

Project Three Rubric –PART ONE

Are questions and observations documented in the report?

Questions and observations are placed regularly throughout the report.

- For each exploration category (univariate, bivariate, multivariate) state the assumptions and questions the charts should answer.
- After each plot or set of related plots, describe what you found.

Tip: Use the ""Question-Visualization-Observations"" framework throughout the exploration.

Tip: For the Part I notebook, use *File* > *Download as...* > *HTML or PDF* menu option to generate the HTML/PDF.

Project Three Rubric –PART Two

Explanatory Data Analysis

Criteria	Submission Requirements
Does the explanatory data analysis tell a story?	The explanatory analysis follows a logical flow and is organized in the following manner: 1. Student presents a clear overarching question or theme that was used to guide their data analysis. 2. Student provides some context for the dataset analyzed. 3. Student presents their findings and key insights. 4. Visualizations from their exploratory analysis are used to support findings. 5. Visualizations should be polished with clear axis, labels, and annotations.

Palestine Launchpad
منصة فلسطين التقنية
Google

Count vs Bar plot Analysis

Charts Types

UDACITY

[Count plots and bar plots | Python (datacamp.com)]

Palestine Launchpad
منصة فلسطين التقنية
Google

Bivariate Analysis

Charts Types

UDACITY

Bivariate Exploration

1. Line Chart
2. Scatterplots - required
3. Box Plots - required
4. Clustered Bar Chart
5. Heatmap

1.Line Chart

Line Plot	https://colab.research.google.com/drive/10l_OoFx8SGL_JVKdStQtLAzNZBJeiP-p?usp=sharing
Exercise	https://colab.research.google.com/drive/1K_lWcYo7SCSpg-qc-c6u4voLqfXpb6ch?usp=sharing
Solution	

Histogram in Seaborn

- https://colab.research.google.com/drive/1nJy6Us1zOOH6Qr46f2enhBCV-deALUG0?usp=sharing

Palestine Launchpad
منصة فلسطين التقنية
Google

New Concepts

Udacity, Students and Session Leads

UDACITY

SELENIUM WEBDRIVER

https://selenium-python.readthedocs.io/

1. Install selenium library
2. Install WebDriver suitable for your browser

Palestine Launchpad
منصة فلسطين التقنية
Google

Knowledge Sahring

Udacity, Students and Session Leads

UDACITY

Text processing pipeline by Sondos

Palestine Launchpad
منصة فلسطين التقنية
Google

Next Steps

Let's get things started...

UDACITY

Second Project July 2, 2024

May June July August September

Strat working on Project three

1. Finish Part 1: Exploration
2. Start with part 2: Explanation

Today
July 6, 2024

Third Project Due Date
August 27, 2024
- Communicate Data findings

Program Period
May 9, 2024 - September 14, 2024

Palestine Launchpad
منصة فلسطين التقنية
Google

Thank You!

And Good Luck!

UDACITY