



REPORT ON INTERNSHIP

UNIFIED MENTOR



INTRODUCTION

This report provides a comprehensive overview of my internship experience at Unified Mentor from February 5 to March 7. The training covered a wide range of topics in data science, including Python programming, data structures, NumPy, Pandas, exploratory data analysis (EDA), statistics, and visualization using Matplotlib, Seaborn, and Tableau. Each day introduced new concepts, hands-on assignments, and practical applications.



Day 1: Project Allocation & Selection

- Introduction to available projects.
- Selection based on expertise and interest.
- Learning modules unlocked with an email notification.

Day 2: Data Science Course Structure & Basics

- Introduction to data science.
- Understanding data types and variables.
- Python operators and their use cases.



Day 3: Python Data Structures

- Lists and their methods (append, remove, sort, etc.).
- Strings and their methods (concatenation, slicing, format()).

Day 4: Conditional Statements, Functions & Loops

- Understanding if-else and elif statements.
- Writing functions and using loops (for and while).

Day 5: Ordered & Unordered Data Structures

- Tuples and their immutability.
- Dictionary operations (keys, values, updating entries).
Sets and their applications.



Day 6: Assignments

- Completion of Python Basics and Data Structure assignments.

Day 7-8: NumPy

- NumPy arrays and dimensions.
- Reshaping, indexing, and slicing.
- Operations on arrays: broadcasting, element-wise operations.

Day 9-11: Pandas for Data Analysis

- Creating and reading DataFrames.
- Indexing, concatenation, and group-by operations.
- Working with date ranges and missing values.



Day 12-13: Exploratory Data Analysis (EDA)

- Handling duplicates and outliers using IQR.
- Extracting missing values and imputation techniques.

Day 14-15: Visualization with Matplotlib & Seaborn

- Plotting basic and advanced charts.
- Seaborn techniques for visualizing data.

Day 16-18: Statistical Analysis

- Measures of central tendency (mean, median, mode).
- Variance and standard deviation.
- Understanding univariate, bivariate, and multivariate analysis.



Day 19-22: Case Studies and Assignments

- Analysis of Wine Quality and Employee datasets.
- Telecom dataset analysis for predictive insights.

Day 23-30: Introduction to Tableau

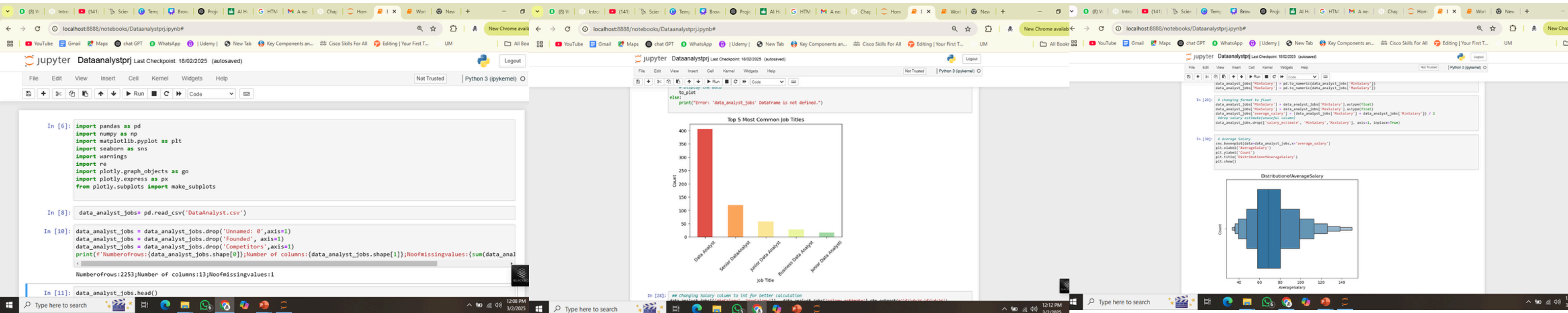
- Learning Tableau's interface and functionalities.
- Working with Superstore and Titanic datasets.
- Creating dashboards and interactive visualizations.

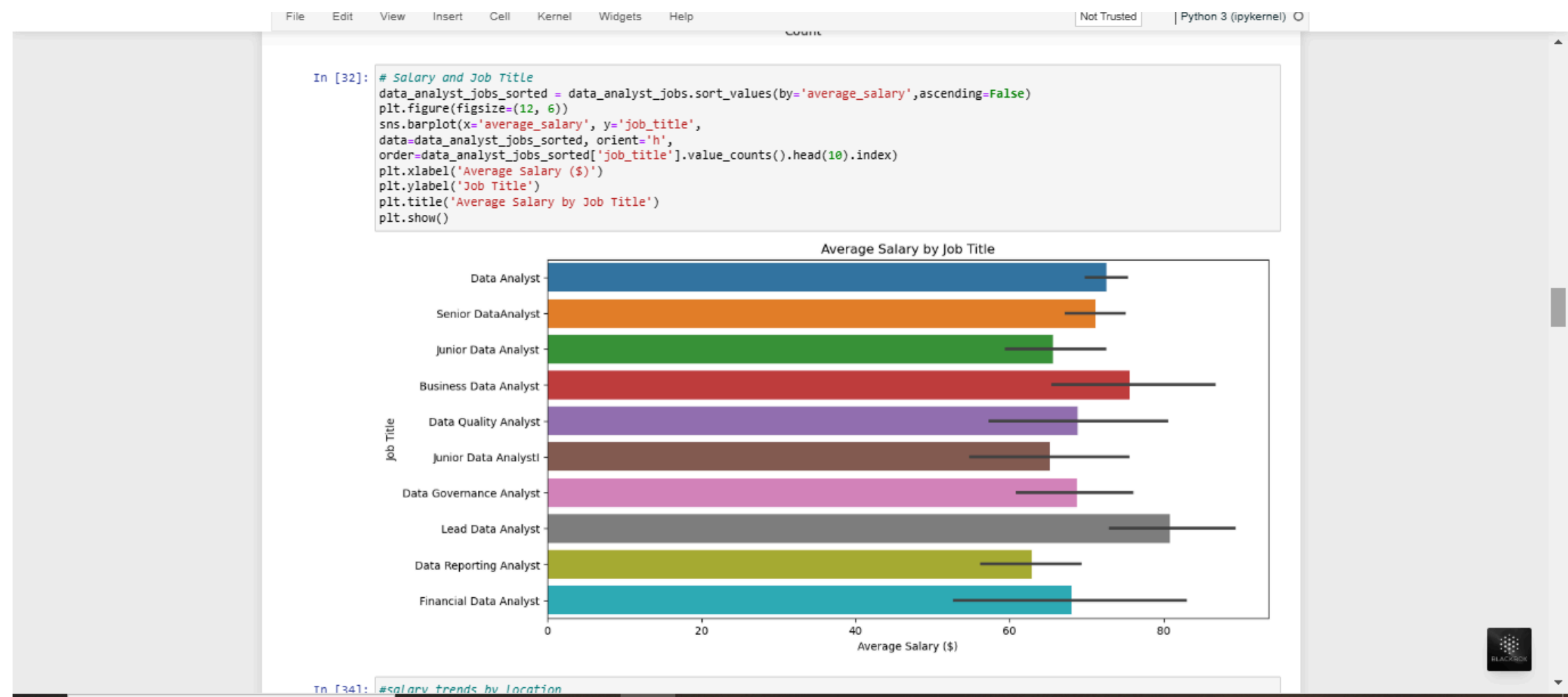
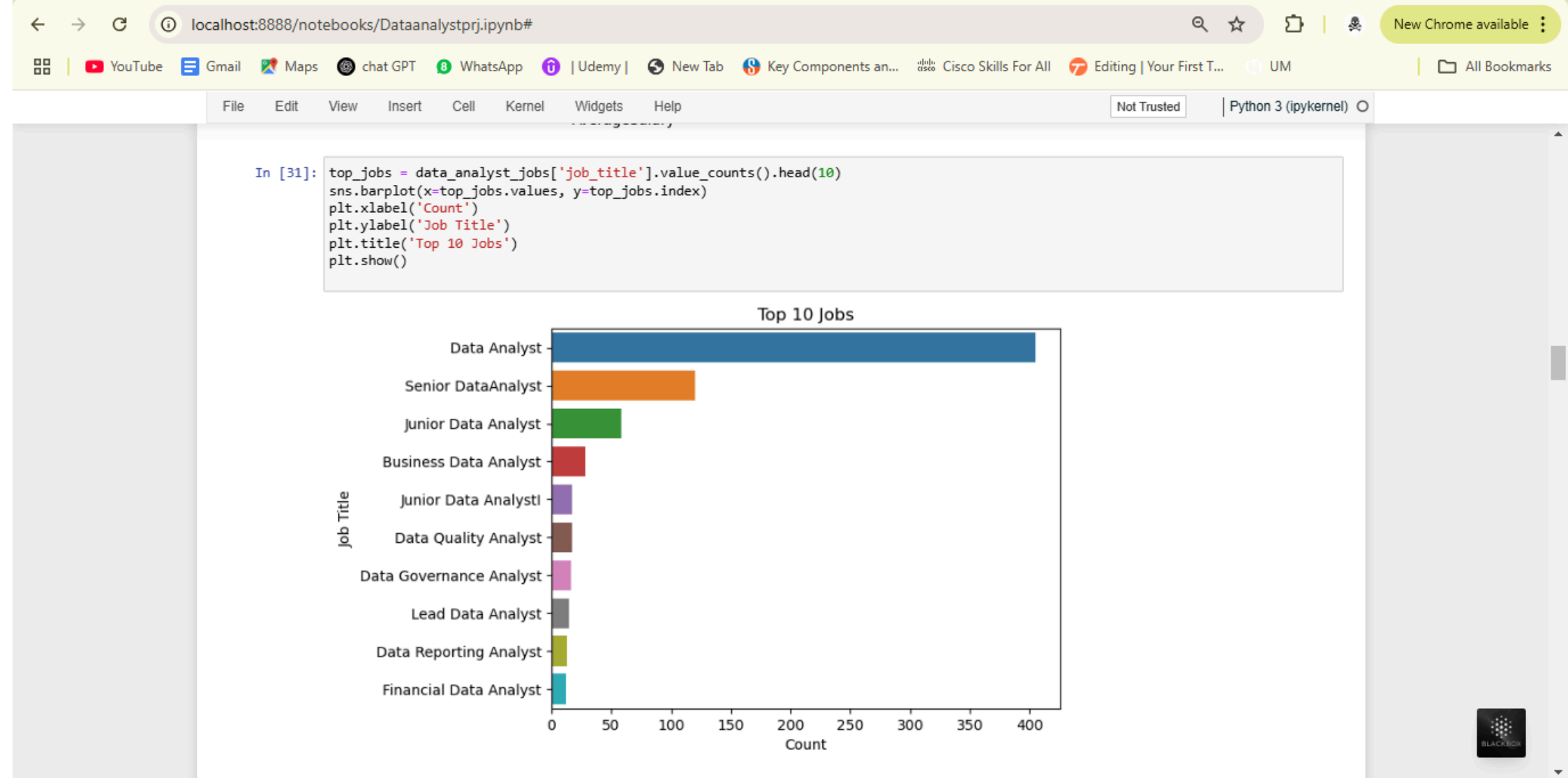
Day 31-32: Final Project - Green Destination Analysis

- Working with the dataset Greendestination.csv.
- Data visualization and insights generation.
- Preparing a final report and PowerPoint presentation.

PROJECT 1: DATA ANALYST JOBS (FINANCE ANALYST)

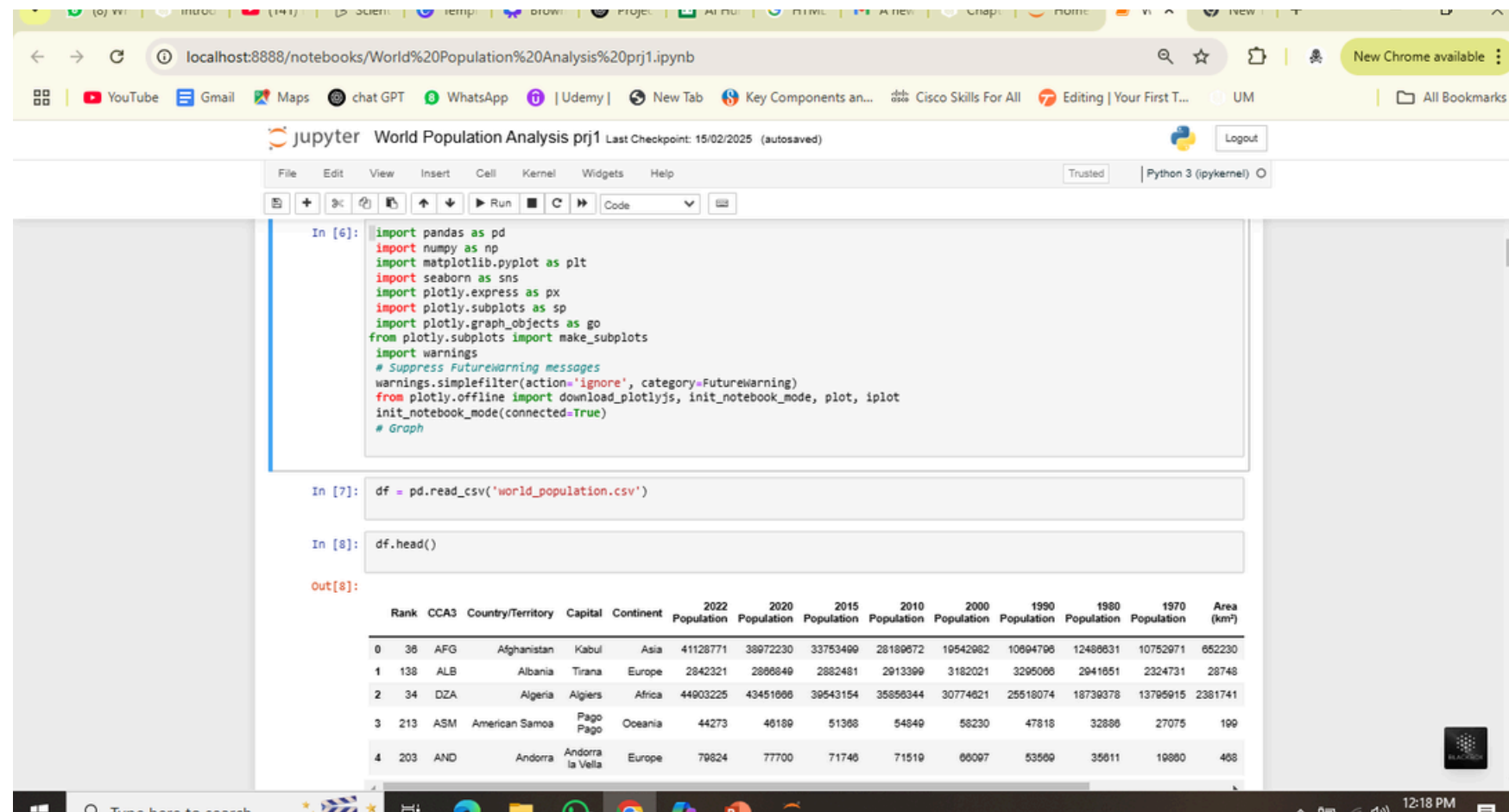
- Analysis of job trends, salary predictions, and skills required.
- Code implementation includes data loading, EDA, visualization, and machine learning models.
- Predicting salary ranges based on company rating and job descriptions.





Project 2: World Population Analysis (Machine Learning)

- Analyzing historical world population data.
- Implementing machine learning techniques for population growth prediction.
- Data visualization using Plotly and Matplotlib.



The screenshot shows a Jupyter Notebook interface with the following code and output:

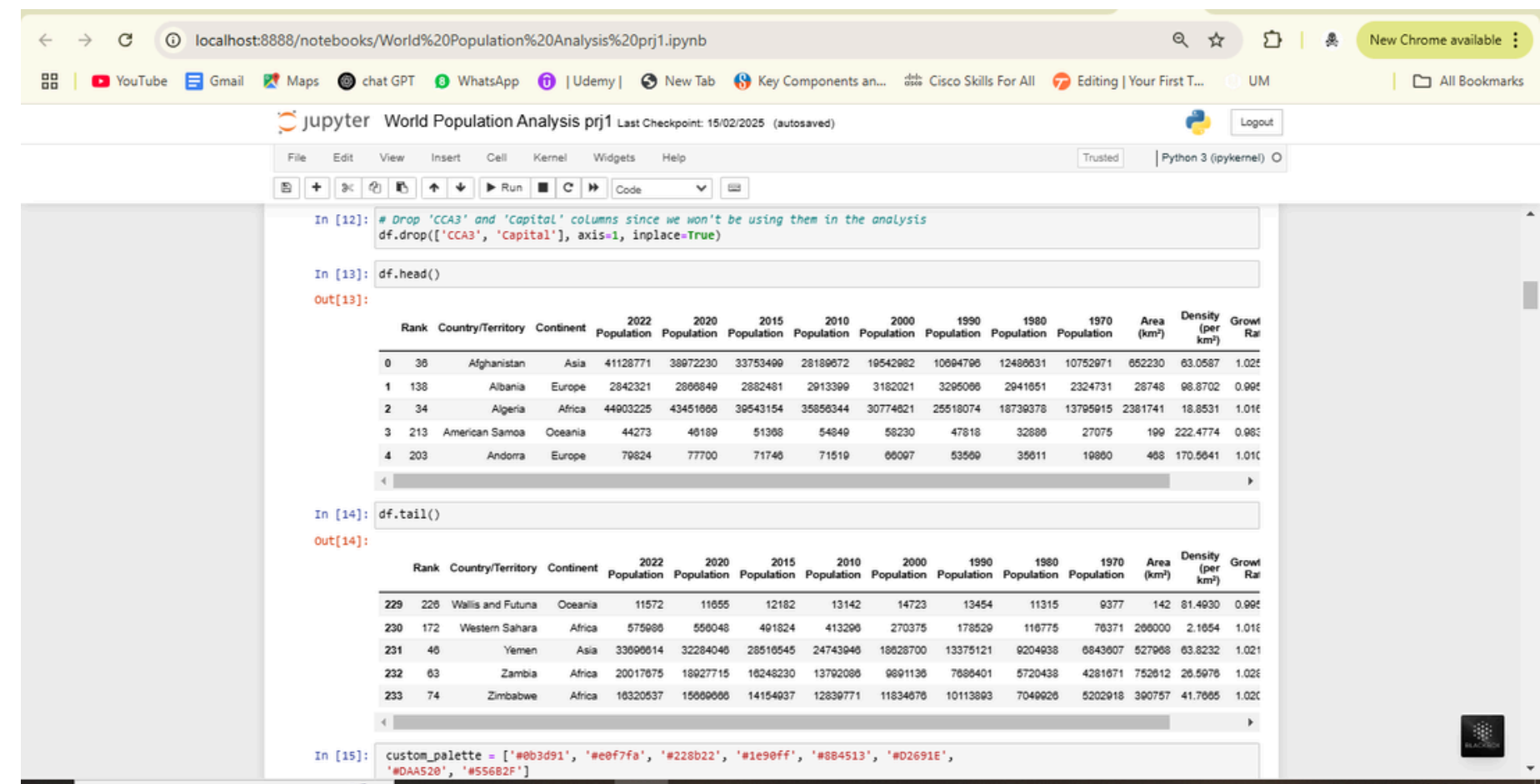
```
In [6]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.subplots as sp
import plotly.graph_objects as go
from plotly.subplots import make_subplots
import warnings
# Suppress FutureWarning messages
warnings.simplefilter(action='ignore', category=FutureWarning)
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
init_notebook_mode(connected=True)
# Graph

In [7]: df = pd.read_csv('world_population.csv')

In [8]: df.head()
```

Out[8]:

	Rank	CCA3	Country/Territory	Capital	Continent	2022 Population	2020 Population	2015 Population	2010 Population	2000 Population	1990 Population	1980 Population	1970 Population	Area (km²)	Density (per km²)	Growth Rate
0	36	AFG	Afghanistan	Kabul	Asia	41128771	38972230	33753499	28189672	19542982	10894796	12489631	10752971	652230	63.0587	1.02%
1	138	ALB	Albania	Tirana	Europe	2842321	2868649	2882481	2913399	3182021	3295086	2941851	2324731	28748	98.8702	0.99%
2	34	DZA	Algeria	Algiers	Africa	44903225	43451686	39543154	35856344	30774821	25518074	18739378	13795915	2381741	18.8531	1.01%
3	213	ASM	American Samoa	Pago Pago	Oceania	44273	46189	51368	54849	58230	47818	32886	27075	199	222.4774	0.98%
4	203	AND	Andorra	Andorra la Vella	Europe	79824	77700	71746	71519	66097	53599	35611	19880	468	170.5641	1.01%



The screenshot shows a Jupyter Notebook interface with the following code and output:

```
In [12]: # Drop 'CCA3' and 'Capital' columns since we won't be using them in the analysis
df.drop(['CCA3', 'Capital'], axis=1, inplace=True)

In [13]: df.head()

Out[13]:
```

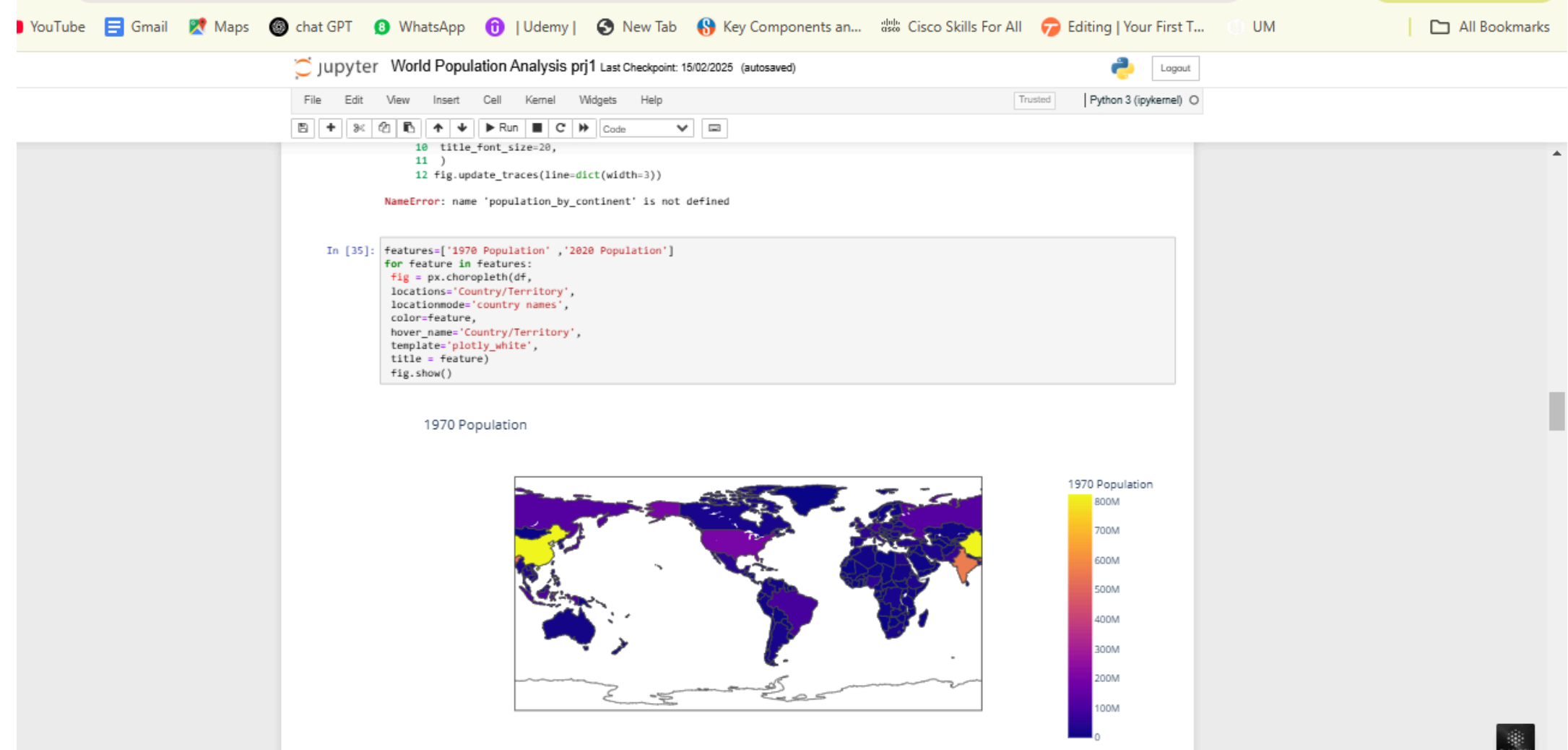
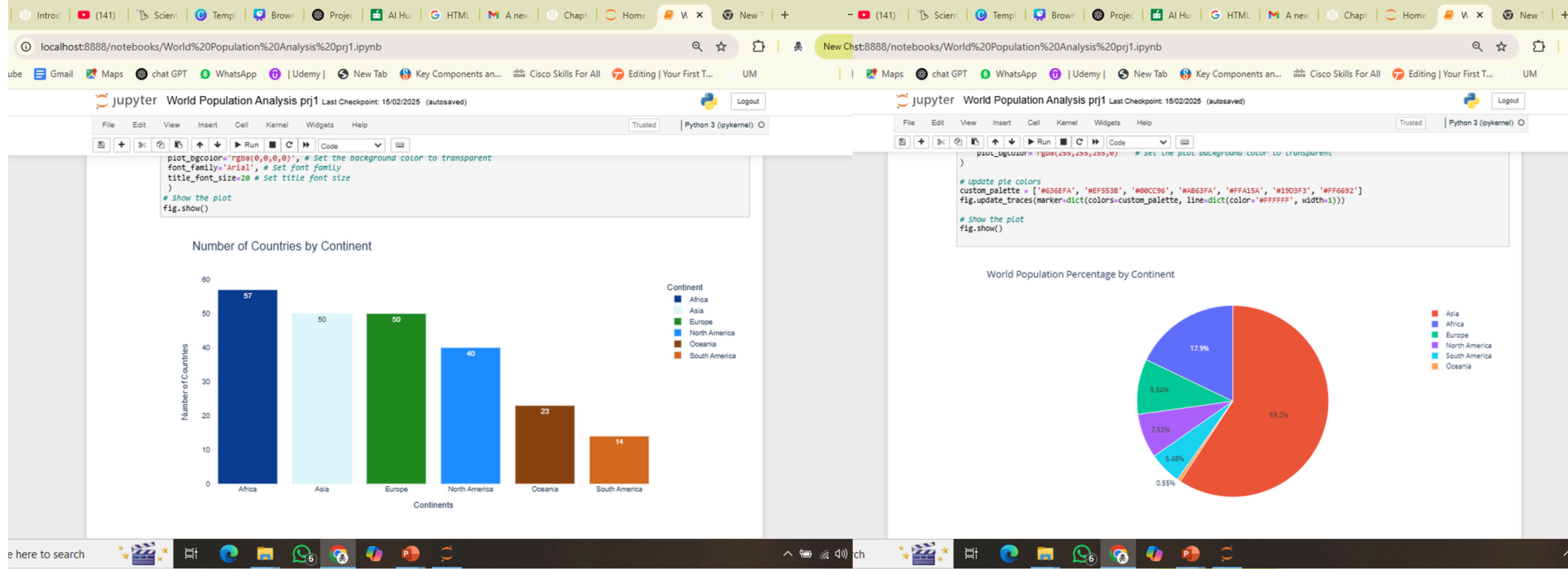
	Rank	Country/Territory	Continent	2022 Population	2020 Population	2015 Population	2010 Population	2000 Population	1990 Population	1980 Population	1970 Population	Area (km²)	Density (per km²)	Growth Rate
0	36	Afghanistan	Asia	41128771	38972230	33753499	28189672	19542982	10894796	12489631	10752971	652230	63.0587	1.02%
1	138	Albania	Europe	2842321	2868649	2882481	2913399	3182021	3295086	2941851	2324731	28748	98.8702	0.99%
2	34	Algeria	Africa	44903225	43451686	39543154	35856344	30774821	25518074	18739378	13795915	2381741	18.8531	1.01%
3	213	American Samoa	Oceania	44273	46189	51368	54849	58230	47818	32886	27075	199	222.4774	0.98%
4	203	Andorra	Europe	79824	77700	71746	71519	66097	53599	35611	19880	468	170.5641	1.01%

```
In [14]: df.tail()

Out[14]:
```

	Rank	Country/Territory	Continent	2022 Population	2020 Population	2015 Population	2010 Population	2000 Population	1990 Population	1980 Population	1970 Population	Area (km²)	Density (per km²)	Growth Rate
229	226	Wallis and Futuna	Oceania	11572	11655	12182	13142	14723	13454	11315	9377	142	81.4930	0.96%
230	172	Western Sahara	Africa	575989	556048	491824	413296	270375	178529	116775	76371	266000	2.1654	1.01%
231	46	Yemen	Asia	33699614	32284046	28518545	24743946	18928700	13375121	9204938	6843607	527968	63.8232	1.02%
232	63	Zambia	Africa	20017675	18927715	16248230	13792086	9691136	7686401	5720438	4281671	752912	26.5978	1.02%
233	74	Zimbabwe	Africa	16320537	15069686	14154937	12839771	11834676	10113893	7049926	5202918	390757	41.7665	1.02%

```
In [15]: custom_palette = ['#0b3d91', '#e0f7fa', '#228b22', '#1e90ff', '#8b4513', '#d2691e',
                        '#daa520', '#556b2f']
```





CONCLUSION

The internship at Unified Mentor provided hands-on experience in data science, covering essential tools and techniques. By working on real-world datasets and implementing machine learning models, I gained practical insights into data manipulation, analysis, and visualization.

This report, along with the attached codes, demonstrates the learning journey and the skills acquired during this internship.

The image features a light cream background with the words "THANK YOU" centered in a bold, brown, sans-serif font. The corners are decorated with stylized, abstract shapes in shades of brown and orange, resembling splatters or organic forms, some with small white dots.

THANK YOU