

★ Member-only story

Storytelling with Exploratory Data Analysis (EDA) in Python



Python Fundamentals · [Follow](#)

4 min read · Jun 4, 2024



20



In the realm of data science, the ability to tell a compelling story with data is as crucial as the technical skills to analyze it. Exploratory Data Analysis (EDA) serves as the backbone of these narratives, enabling data scientists to uncover patterns, relationships, and insights. In this article, we'll explore the art of storytelling with data using Python for EDA. Through a series of examples and practical code snippets, we'll unravel the techniques that transform raw data into engaging narratives, making your insights accessible and impactful.



Photo from [Pexels](#)

Chapter 1: Setting the Stage

Before we delve into the code, let's set the stage for our narrative. Load your dataset and conduct an initial exploration:

```
import pandas as pd

# Load your dataset
data = pd.read_csv('your_dataset.csv')

# Display the first few rows
print(data.head())
```

Chapter 2: Unveiling the Story through Descriptive Statistics

EDA begins with understanding the basic characteristics of your data. Use descriptive statistics to reveal the story:

```
# Descriptive statistics
print(data.describe())
```

Chapter 3: The Missing Pieces: Handling Missing Values

Every narrative has its missing pieces. Identify and address them in your data:

```
# Check for missing values
print(data.isnull().sum())

# Handle missing values (if needed)
# data = data.dropna() # or impute missing values
```

Chapter 4: Data Visualization the Journey: Plots and Graphs

Visualizations are the heart of data narratives. Create compelling plots to illustrate key points:

Section 1: Introduction to our Example Dataset

```
import pandas as pd
import matplotlib.pyplot as plt

# Creating a basic example dataset
data = {
```

```
'Month': ['Jan', 'Feb', 'Mar', 'Apr', 'May'],
'Sales': [150, 200, 180, 220, 250],
'Expenses': [80, 100, 90, 110, 120]
}

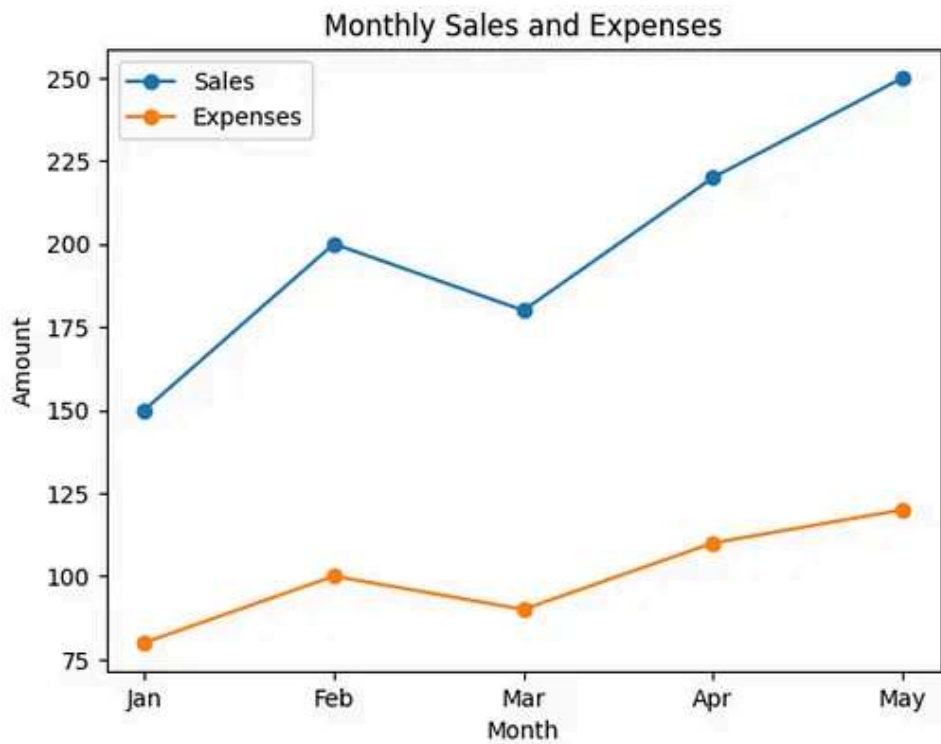
df = pd.DataFrame(data)
print(df)
```

	Month	Sales	Expenses
0	Jan	150	80
1	Feb	200	100
2	Mar	180	90
3	Apr	220	110
4	May	250	120

Section 2: Line Plots — The Classic Visualization

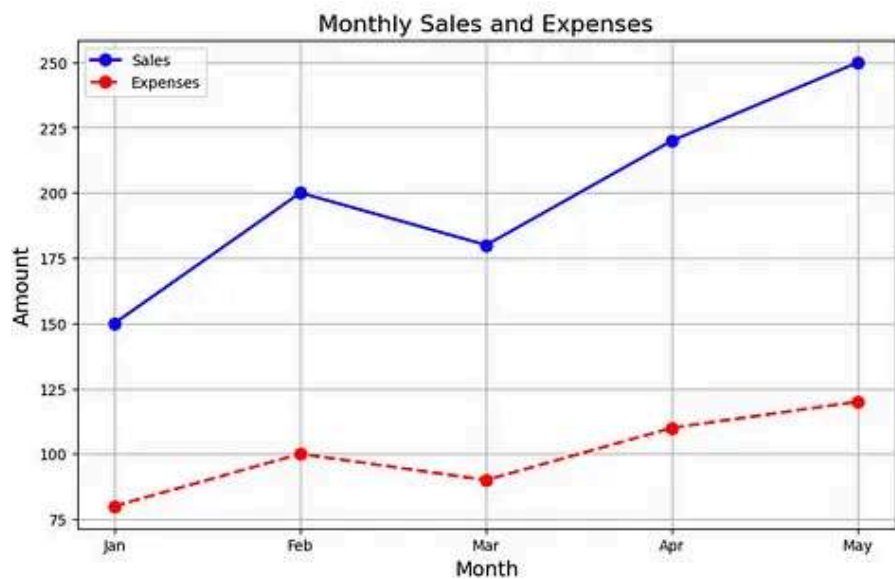
2.1 Simple Line Plot

```
# Simple line plot
plt.plot(df['Month'], df['Sales'], marker='o', label='Sales')
plt.plot(df['Month'], df['Expenses'], marker='o', label='Expenses')
plt.xlabel('Month')
plt.ylabel('Amount')
plt.title('Monthly Sales and Expenses')
plt.legend()
plt.show()
```



2.2 Customized Line Plot

```
# Customized line plot
plt.figure(figsize=(10, 6))
plt.plot(df['Month'], df['Sales'], marker='o', color='blue', linestyle='-', line
plt.plot(df['Month'], df['Expenses'], marker='o', color='red', linestyle='--', l
plt.xlabel('Month', fontsize=14)
plt.ylabel('Amount', fontsize=14)
plt.title('Monthly Sales and Expenses', fontsize=16)
plt.legend()
plt.grid(True)
plt.show()
```



Section 3: Bar Plots — Effective for Comparisons

3.1 Basic Bar Chart

```
# Basic bar chart
plt.bar(df['Month'], df['Sales'], color='skyblue', label='Sales')
plt.bar(df['Month'], df['Expenses'], color='lightcoral', label='Expenses', bottom=True)
plt.xlabel('Month')
plt.ylabel('Amount')
plt.title('Monthly Sales and Expenses (Bar Chart)')
plt.legend()
plt.show()
```

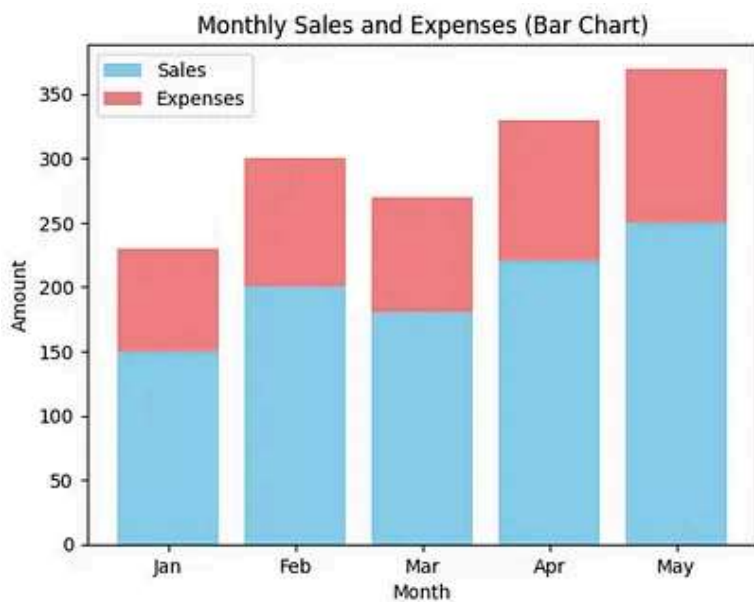
[Open in app](#)

Medium

 Search

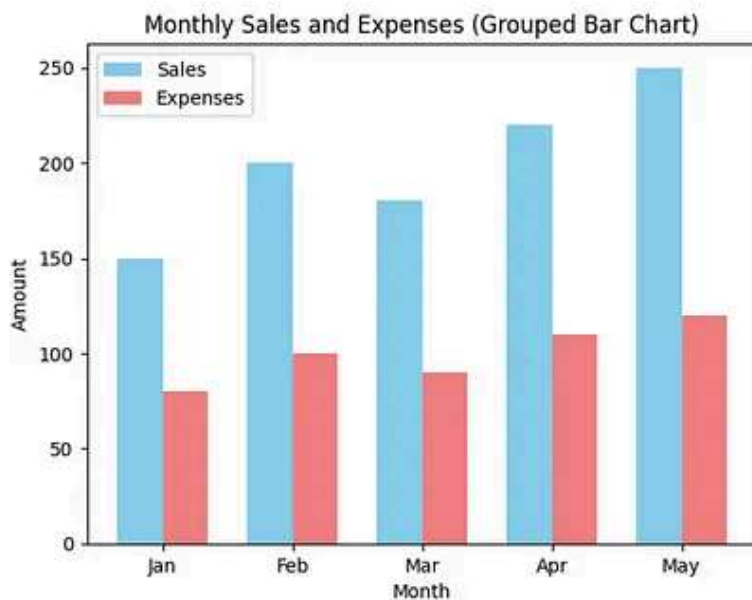
 Write





3.2 Grouped Bar Chart

```
# Grouped bar chart
bar_width = 0.35
bar_positions = range(len(df['Month']))
plt.bar(bar_positions, df['Sales'], width=bar_width, color='skyblue', label='Sales')
plt.bar([pos + bar_width for pos in bar_positions], df['Expenses'], width=bar_width, color='red', label='Expenses')
plt.xlabel('Month')
plt.ylabel('Amount')
plt.title('Monthly Sales and Expenses (Grouped Bar Chart)')
plt.xticks([pos + bar_width / 2 for pos in bar_positions], df['Month'])
plt.legend()
plt.show()
```



Chapter 5: The Data Arc: GroupBy Operations

Bring the narrative full circle by leveraging GroupBy operations for in-depth analysis:

```
# GroupBy operations
average_age_by_category = data.groupby('category_column')['Age'].mean()
print(average_age_by_category)
```

Conclusion:

In the realm of data science, the ability to tell a story with data is an art. Through effective Exploratory Data Analysis in Python, you can craft compelling narratives that captivate your audience. As you embark on your storytelling journey, remember that each dataset has its unique story waiting to be uncovered. Adapt these techniques, experiment with different visualizations, and let the data guide you in weaving a narrative that not only informs but also resonates.

Python Fundamentals

Thank you for your time and interest! 🚀

You can find even more content at [Python Fundamentals](#) 🌟

Python's Gurus 🚀

Thank you for being a part of the [Python's Gurus community](#)!

Before you go:

- Be sure to **clap x50 time** and **follow** the writer 🙌
- Follow us: [Newsletter](#)
- Do you aspire to become a Guru too? *Submit your best article or draft* to reach our audience.

Python

Data Science

Data Analysis

Data Scientist

Data Analyst



Written by Python Fundamentals


2K Followers

Follow



More from Python Fundamentals



 Python Fundamentals in Python in Plain English

15 Advanced Code Snippets for Python Data Scientists and Data...


As a data scientist or data analyst, mastering advanced coding techniques is essential for...

🌟 · 6 min read · May 18, 2024



22



 Python Fundamentals

Data Visualization for Exploratory Data Analysis (EDA) in Python

Python Data Visualization Guide

🌟 · 6 min read · Feb 25, 2024



544



2



 Python Fundamentals in Python's Gurus

Python Data Manipulation: Sorting, Grouping, Aggregation, Merging

Data Manipulation with Python

🌟 · 5 min read · May 26, 2024



 Python Fundamentals in Towards Dev

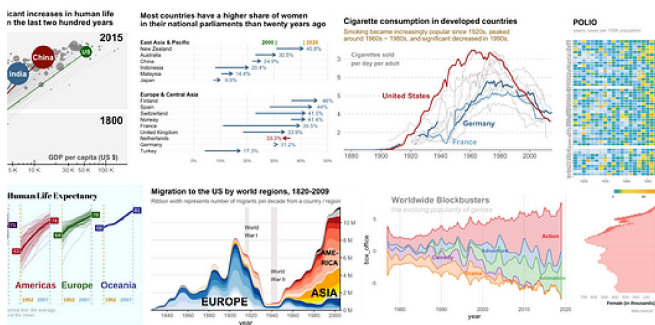
20 Python Pandas Codes To Elevate Your Data Analysis Skills

Pandas, the powerhouse library for data manipulation and analysis in Python, offers a...

🌟 · 3 min read · May 22, 2024

[See all from Python Fundamentals](#)

Recommended from Medium

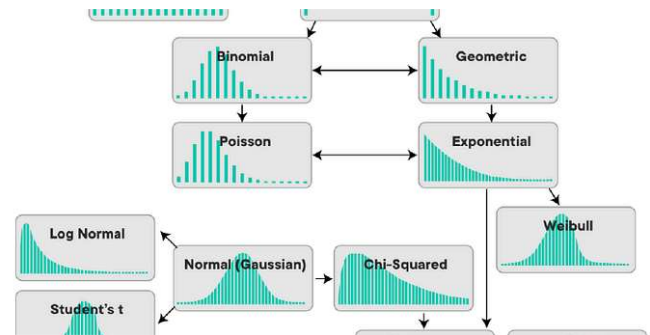


Bo Yuan, Ph.D.

Awesome Strategies to Visualize Change with Time

This article summarizes effective strategies to visualize temporal changes, illustrated with...

5 min read · May 8, 2024



Sachinsoni

Understanding Different Kinds of Distributions in Statistics

Statistics is a powerful tool for making sense of data, and at its core lies the concept of...

8 min read · Apr 26, 2024



Predictive Modeling w/ Python

20 stories · 1288 saves



Practical Guides to Machine Learning

10 stories · 1548 saves



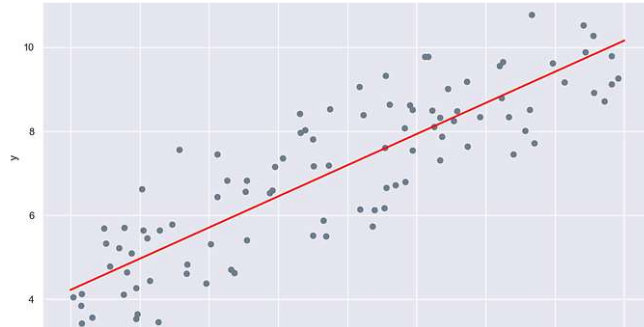
Coding & Development

11 stories · 652 saves



ChatGPT prompts

48 stories · 1671 saves



 Rakesh Ganya

From Data to Predictions: Understanding Linear Regression

Linear Regression

7 min read · Jun 6, 2024



 Miriam Santos in Towards Data Science

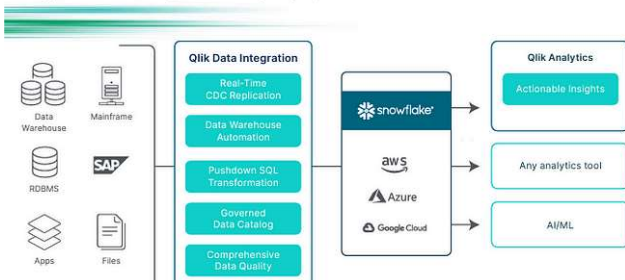
A Data Scientist's Essential Guide to Exploratory Data Analysis

Best Practices, Techniques, and Tools to Fully Understand Your Data

11 min read · May 30, 2023



Together, Qlik and Snowflake offer an end-to-end, data-centric solution – topped off with Talend's pervasive data quality capabilities.



 Kavya in DataDrivenInvestor

9 Best Data Visualization Tools in 2024



 Shana Nasrin in Learning Data

Data Analyst Portfolio Project (Excel)

A Comparison on Key features, and best use cases

🌟 · 7 min read · Jun 3, 2024


 48 

Airline Passenger Satisfaction

7 min read · Apr 11, 2024

 142  4

See more recommendations