Python Programming Language Logo

Introduction to Data Science with Python

Python has become the leading language for data science and machine learning applications.

- Easy to learn and use
- Rich ecosystem of libraries
- Strong community support

Essential Python Libraries

Data Analysis

NumPy - Numerical computing
Pandas - Data manipulation
SciPy - Scientific computing

[Table content follows]

Library | Primary Use | Key Features
NumPy | Numerical Computing | Arrays, Broadcasting, Linear Algebra
Pandas | Data Manipulation | DataFrames, Time Series, I/O Tools
Matplotlib | Visualization | Plots, Charts, Customization

Data Visualization

Creating compelling visualizations

Common visualization types in data science

[Interactive Chart: Monthly Sales Data]

Python Code Examples

Basic Data Analysis with Pandas

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

# Load and explore data
df = pd.read_csv('sales_data.csv')
print(df.head())
```

Output of the above code

```
# Basic statistics
print(df.describe())

# Visualization
df.groupby('month')['sales'].sum().plot(kind='bar')
plt.title('Monthly Sales')
plt.xlabel('Month')
plt.ylabel('Sales ($)')
plt.show()
```

Machine Learning Pipeline

1. Data Collection - Gathering relevant data from various sources
2. Data Preprocessing - Cleaning and transforming raw data
3. Feature Engineering - Creating meaningful features for models
4. Model Training - Building and tuning ML algorithms
5. Evaluation - Assessing model performance
6. Deployment - Implementing the model in production

Case Study: Customer Segmentation

Challenge
A retail company wants to segment its customers based on purchasing behavior to create targeted marketing campaigns.
Approach

RFM Analysis (Recency, Frequency, Monetary)
K-means Clustering
[Interactive Scatter Plot: Customer Segments]

Q&A Session

Contact Information

Email: presenter@example.com
Twitter: @datascientist
GitHub: github.com/datascientist

Download slides and code examples: example.com/resources