Python

Seaborn

Alish Bista - 11/11/2025

Setting Up VSCode

venv

virtual environment

- Isolated Python workspace for each project
- Prevents package conflicts between projects
- Makes your work reproducible and shareable

```
python3 -m venv .venv
source .venv/bin/activate
pip install seaborn pandas matplotlib
```

```
python -m venv .venv
.venv\Scripts\activate.ps1
pip install seaborn pandas matplotlib
```

Seaborn

Introduction to Seaborn

- Built on top of Matplotlib for beautiful graphics
- Works seamlessly with pandas DataFrames
- Built-in themes and color palettes
- Specialized plots for statistical analysis

```
import seaborn as sns
import matplotlib.pyplot as plt

# Load a dataset
tips = sns.load_dataset('tips')
print(tips.head())

# See all available datasets
print(sns.get_dataset_names())
```

relplot

Relationship Plots - Scatter

- Creates scatter plots by default (kind='scatter')
- Switch to line plots with kind='line'
- Easy faceting with col and row parameters
- Handles multiple dimensions with hue, size, style

```
# Scatter plot with categories

sns.relplot(data=tips, x='total_bill',
y='tip', hue='time', size='size',
style='smoker')

plt.show()
```

```
# Faceted scatter plots

sns.relplot(data=tips, x='total_bill',
y='tip', col='time', row='sex')

plt.show()
```

relplot

Relationship Plots - Line

- Set kind='line' for line plots
- Automatic aggregation with confidence intervals
- Great for comparing trends across categories

```
flights = sns.load_dataset('flights')

# Single line plot
sns.relplot(data=flights, x='year', y='passengers',
            kind='line')
plt.show()

# Multiple lines with facets
sns.relplot(data=flights, x='year', y='passengers',
            hue='month', kind='line', col='month',
col_wrap=4)
plt.show()
```

catplot

Categorical Plots

- Multiple plot kinds: box, violin, bar, point
- Automatically chooses appropriate plot type
- Built-in faceting support
- Perfect for comparing groups

```
# Strip plot (default)
sns.catplot(data=tips, x='day', y='total_bill')
plt.show()

# Box plot
sns.catplot(data=tips, x='day', y='total_bill',
                    kind='box', hue='time')
plt.show()
```

catplot

Categorical Plots - Distribution in Categories

- kind='box' - quartiles and outliers
- kind='violin' - full distribution shape

```
# Violin plot
sns.catplot(data=tips, x='day', y='total_bill',
                kind='violin', hue='sex', split=True)
plt.show()


# Faceted box plots
sns.catplot(data=tips, x='day', y='total_bill',
                kind='box', col='time')
plt.show()
```

catplot

Categorical Plots - Point Distribution

- kind='strip' - jittered points
- kind='swarm' - non-overlapping points
- Useful for smaller datasets

```
# Strip plot with jitter

sns.catplot(data=tips, x='day',
y='total_bill', kind='strip',
hue='time')

plt.show()


# Swarm plot (no overlap)

sns.catplot(data=tips, x='day',
y='total_bill', kind='swarm',
hue='smoker')

plt.show()
```

Distribution Plots

Histogram & KDE

- histplot() - binned frequency counts
- kdeplot() - smooth density estimate
- Both support univariate and bivariate plots

```
# Histogram with KDE overlay
sns.histplot(data=tips, x='total_bill', kde=True)
plt.show()

# Multiple distributions
sns.kdeplot(data=tips, x='total_bill', hue='time', fill=True)
plt.show()


# 2D distribution
sns.kdeplot(data=tips, x='total_bill', y='tip', fill=True)
plt.show()
```

Heat Maps

Plotting The Heat Matrix

- Shows data values as colors in a grid
- Perfect for correlation matrices
- annot=True displays values in cells
- Customizable color palettes

```
# Correlation heatmap
numeric_tips = tips.select_dtypes(include='number')
corr = numeric_tips.corr()


sns.heatmap(corr, annot=True, cmap='coolwarm',
center=0)
plt.show()


# Custom heatmap
flights_pivot = flights.pivot(index='month',
columns='year', values='passengers')
sns.heatmap(flights_pivot, cmap='YlOrRd')
plt.show()
```

Pair Plots

All variables relationships

- Scatter plots for all variable pairs
- Distributions on diagonal
- Color by categories with hue
- Great for initial data exploration

```
iris = sns.load_dataset('iris')


# Basic pairplot

sns.pairplot(iris)

plt.show()


# With categorical coloring

sns.pairplot(iris, hue='species',
diag_kind='kde')

plt.show()
```

Customization

Styling The Charts

- Built-in themes: darkgrid, whitegrid, dark, white, ticks
- Color palettes: deep, muted, pastel, dark, colorblind
- Context settings: paper, notebook, talk, poster
- Easy customization with set_theme()

```
# Set style and palette
sns.set_theme(style='whitegrid', palette='muted')

# Set context for presentation
sns.set_context('talk')

# Custom palette
custom_colors = ['#FF6B6B', '#4ECDC4', '#45B7D1']
sns.set_palette(custom_colors)

sns.catplot(data=tips, x='day', y='total_bill',
kind='box')
plt.show()
```

Q&A