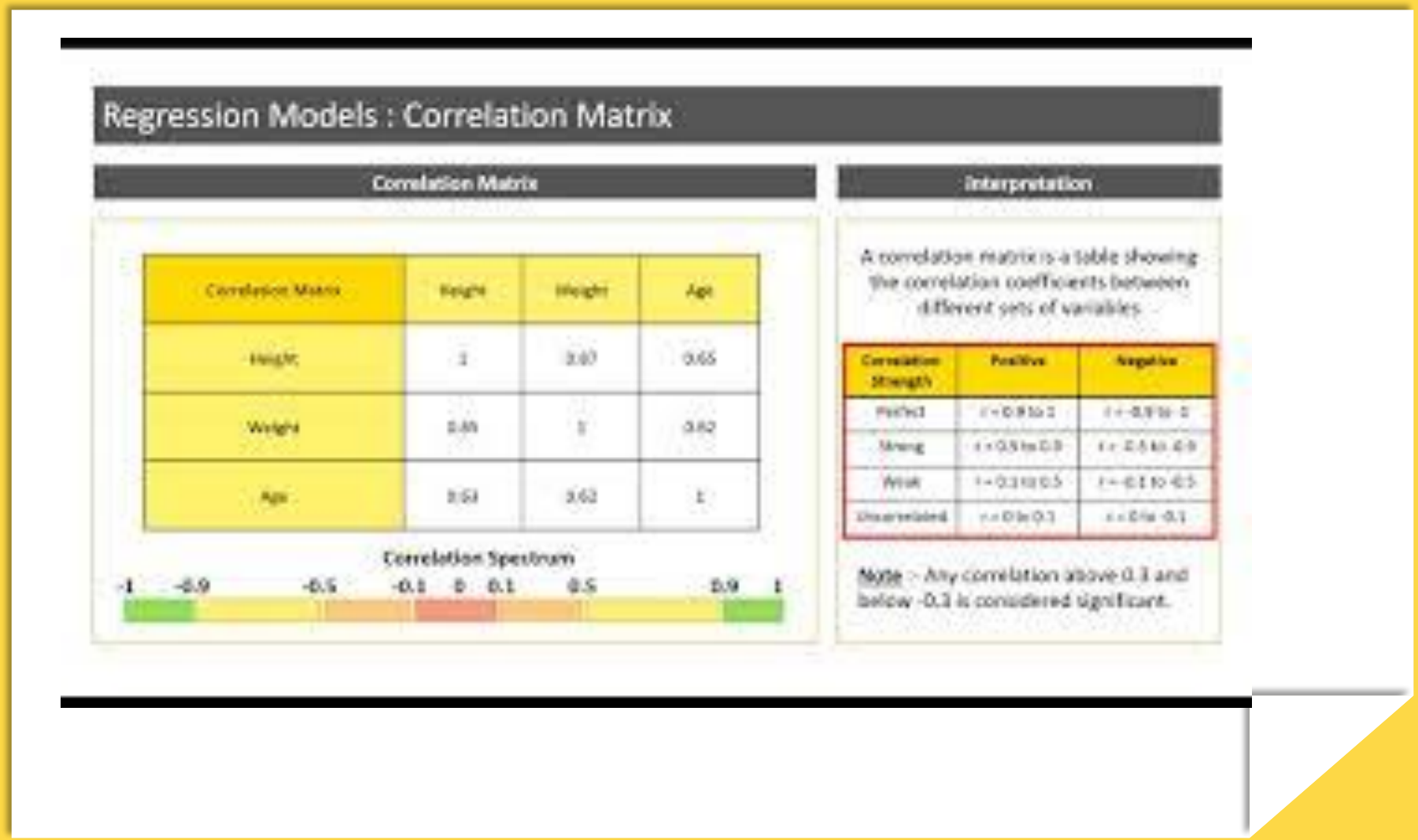
Correlation Matrices Video



Introduction to Correlation Matrices and Heatmaps

Understanding Variable Relationships:

- Variables in a dataset often interact and influence each other.

Introduction to Correlation Matrices:

- Correlation measures the strength and direction of relationships.
- Values range from -1 (strong negative correlation) to 1 (strong positive correlation).

Value of Heatmaps in Visualizing Correlation Matrices:

- Heatmaps provide a visual representation of correlation matrices.
- Color gradients indicate the magnitude of correlations.

Understanding Correlation & Correlation Matrices

Correlation: Measuring Relationship Strength

- Correlation quantifies the degree of association between variables.
- Indicates how changes in one variable relate to changes in another.

Correlation Matrices: Exploring Pairwise Relationships

- Correlation matrices present relationships between all variable pairs.
- Rows and columns represent variables, and cells contain correlation values.

Types of Correlation: Positive, Negative, None

- Positive correlation: Both variables increase or decrease together.
- Negative correlation: One variable increases as the other decreases.
- No correlation: Variables show minimal or no distinct relationship.

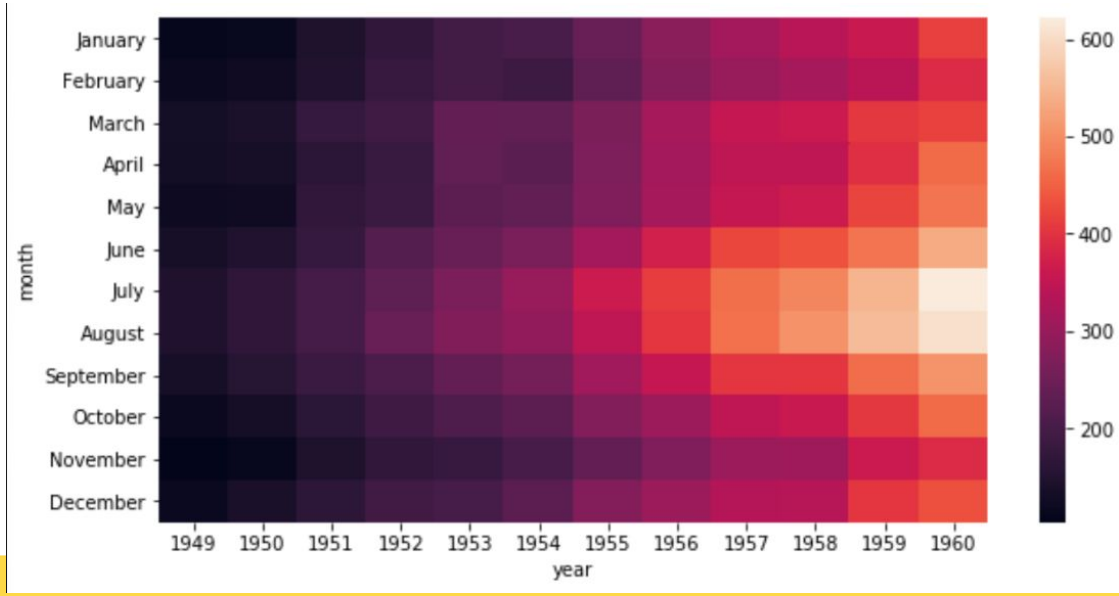
Introduction to Seaborn's Heatmap Fn

Seaborn's sns.heatmap() Function

- Creates heatmap visualizations effortlessly.
- Offers a powerful tool for visualizing data relationships.

Revealing Patterns and Relationships

- Heatmaps are ideal for unveiling hidden patterns in datasets.
- Helps identify connections and trends within data.

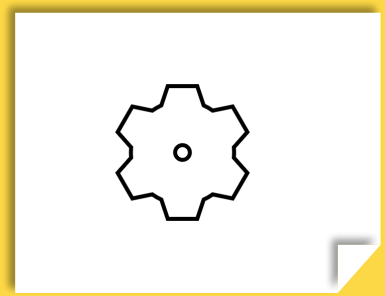


Visualizing Correlation Matrices with Heatmaps

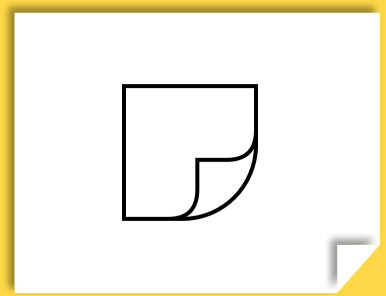
Benefits of Heatmaps:

1) **Quick Identification**: Color variations allow rapid detection of correlation strength.
2) **Intuitive Color Coding**: Colors represent correlations in an easily understandable manner.
3) **Enhanced Interpretation**: Aids in uncovering insights without requiring detailed analysis.

Creating a Correlation Heatmap



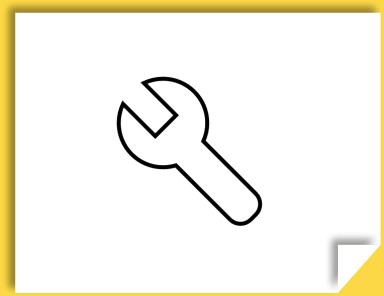
Step 1:
Extract correlation matrix from your data.



Step 2:
Utilize Pandas' .corr() or similar methods.

Step 3:
Generate Heatmap with sns.heatmap():

- Import Seaborn library and your dataset.
- Utilize sns.heatmap() to create the heatmap.
- Pass the correlation matrix to the function.



Step 4:
Customize the Heatmap:

- Use cmap parameter to select color palette.
- Add annotations using annot parameter.
- Enhance visual clarity with labels and titles.

How this would look!

```
import seaborn as sns
import pandas as pd

# Load the example dataset
df = sns.load_dataset("iris")

# Create the correlation matrix
corr_matrix = df.corr()

# Create the heatmap
sns.heatmap(corr_matrix, cmap="coolwarm")
```

Interpreting Heatmaps

1. Reading the Heatmap Color Scale: -1 to 1:

- Color scale represents correlation values.
- Ranges from -1 (strong negative) to 1 (strong positive).

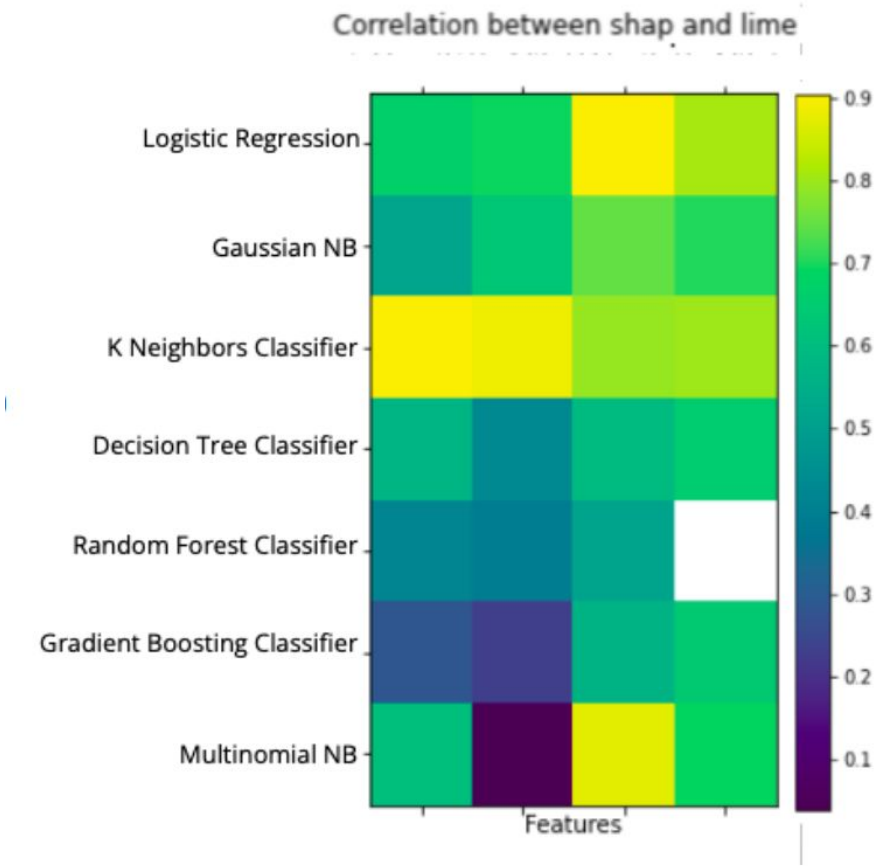
2.Identifying Strong Correlations: Dark and Light Colors:

- Dark colors indicate high correlation.
- Light colors signify weaker correlation or absence.

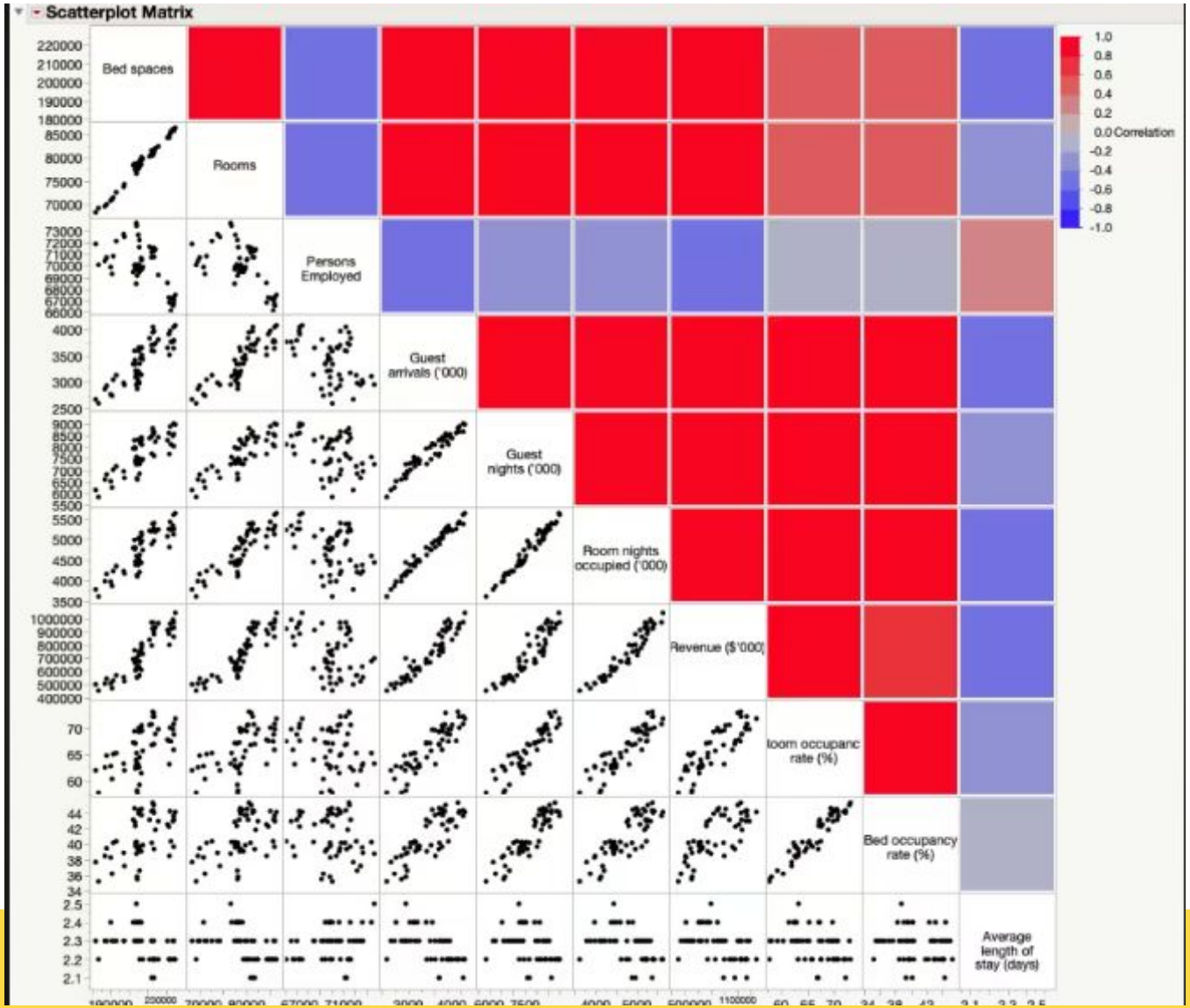
3.Detecting Patterns, Clusters, and Relationships:

- Observe diagonal patterns (self-correlations).
- Identify clusters of similar colors.
- Discern relationships among variables.

Example of Interpreting Heatmaps!



Sample Correlation Matrix



THANK YOU

I hope you enjoyed learning about Heatmaps!