



STRANGER STATS

Data Science Cheat Sheet

Inspecting Data

`df.head(n)` | Shows first n rows of a DataFrame
`df.tail(n)` | Shows last n rows
`df.shape` | Returns (rows, columns)
`df.columns` | List column names
`df.dtypes` | Show data types of columns

Selecting & Filtering

`df['Column']` | Access a single column
`df[['Col1','Col2']]` | Access multiple columns
`df.loc(row_cond, col_cond)` | Select r&c
`df(df['Column'] > value)` | Filter rows
`df.sample(n)` | Random sample of n rows

Important Statistics

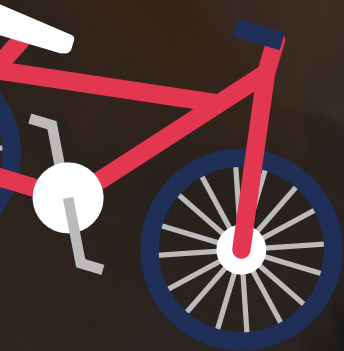
`df['Column'].sum()` | Sum of values
`df['Column'].mean()` | Mean value
`df['Column'].median()` | Median value
`df['Column'].mode()` | Mode value
`df['Column'].var()` | Variance

Visualization

`plt.plot(x, y)` | Line
`plt.bar(x, y)` | Bar chart
`plt.hist(df['Column'])` | Histogram
`plt.show()` | Display plot

Miscellaneous

`Fdf.dropna()` | Drop rows with missing values
`df.fillna(value)` | Fill missing values
`df['Column'].map(func)` | Apply function to each value
`round(df['Column'], 2)` | Round numbers
`df['Column'].dt.hour` | Extract hour from datetime
`df['Column'].dt.minute` | Extract minute from datetime



@css_tamucc

