

Python Implementation and Practice

Key Python Libraries

NumPy

Matrix operations

Scikit-learn

ML convenience functions

Pandas

Data manipulation

Statsmodels

Statistical output

Matplotlib/Seaborn

Visualization

Standard Workflow

1 Load Data



2 Explore & Visualize



3 Fit Model



4 Validate (diagnostics)



5 Predict

Implementation Examples

```
# NumPy (from scratch)
β = np.linalg.inv(X.T @ X) @ X.T @ y

# Scikit-learn (easy)
from sklearn.linear_model import LinearRegression
```

Best Practices

Feature scaling

Handle missing data

```
model = LinearRegression()  
model.fit(X, y)  
  
# Statsmodels (detailed stats)  
import statsmodels.api as sm  
results = sm.OLS(y, X).fit()
```

Train-test split

Cross-validation

Practice Datasets

Boston Housing

California Housing

Diabetes