

```
In [ ]: !pip install -q qbstyles

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

%config InlineBackend.figure_format = 'retina'

from qbstyles import mpl_style
mpl_style(dark=True)

from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import confusion_matrix, classification_report, roc_auc_score

import statsmodels.api as sm

print('imports completed – ready to load data')
```

imports completed – ready to load data

Learning objectives

- Understand how to prepare data for a basic classification problem.
- Build and interpret a logistic regression model.
- Perform model-level (LLR) and coefficient-level hypothesis tests.
- Evaluate model performance and create a simple prediction function.

```
In [ ]: # Load the dataset
url = "https://raw.githubusercontent.com/Kartavya-Jharwal/Kartavya_Business_Analyti
df = pd.read_csv(url)

# Display the first few rows
df.head()
```

```
Out[ ]:
```

	Gender_Coded	Gender	Price	Purchase_Coded	Purchase
0	1	Male	8.49	1	Yes
1	1	Male	12.49	0	No
2	0	Female	11.49	0	No
3	1	Male	9.49	0	No
4	1	Male	14.49	0	No

Automatic target / feature split — prefer 'Upgraded' when present

If 'Upgraded' is present use it as the target; otherwise try to find a binary column (0/1) automatically

Quick EDA — preview, distributions and simple visual checks

Exploratory Data Analysis

Model A: Purchase ~ Price

We will first build a logistic regression model to predict `Purchase` using only `Price`.

Model B: Purchase ~ Price + Gender

Now we will build a second model including `Gender` as a predictor.

Odds Ratios (Model B)

We calculate the odds ratios by taking the exponential of the coefficients.

Predictions

Predict the probability of purchase for a customer when the price is \$10.49, for both Male and Female.