Guide: Importing Online Data & Running ML Models

1. Importing Online Datasets

The most common way to import data from a web URL is using the pandas library. Ensure your file is in a readable format like CSV or Excel.

```
import pandas as pd

# Example: Importing a dataset from a public URL
url = "https://raw.githubusercontent.com/datasciencedojo/datasets/maste
df = pd.read_csv(url)

# Preview the data
print(df.head())
```

2. Preparing Data for Models

Before training, you must define your features (X) and target (y), and handle missing values if necessary.

```
# Select features and target
X = df[['Pclass', 'Age', 'SibSp']] # Replace with actual column names
y = df['Survived']

# Handling missing values
X = X.fillna(X.mean())
```

3. Implementing Models

Use `scikit-learn` to train and evaluate your models.

```
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score

# Split data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.

# Train Random Forest
model = RandomForestClassifier(n_estimators=100)
model.fit(X_train, y_train)

# Evaluate
predictions = model.predict(X_test)
print(f"Accuracy: {accuracy_score(y_test, predictions)}")
```