COURSE / KURS

ReDI School of Digital Integration

Ecommerce Sales Analysis and Forecast

Presenter: Maya

Project summary

Problem statement

E-stores collect a lot of historical data on the sales performance of different products. Forecasting the sales value is required for business and inventory planning, so we will build a machine learning model to forecast these values.

Dataset

The dataset includes transnational data for all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered online retail selling unique all-occasion gifts.

Objective

Analyse historicals sales data and develop the ML model to forecast future sales in units

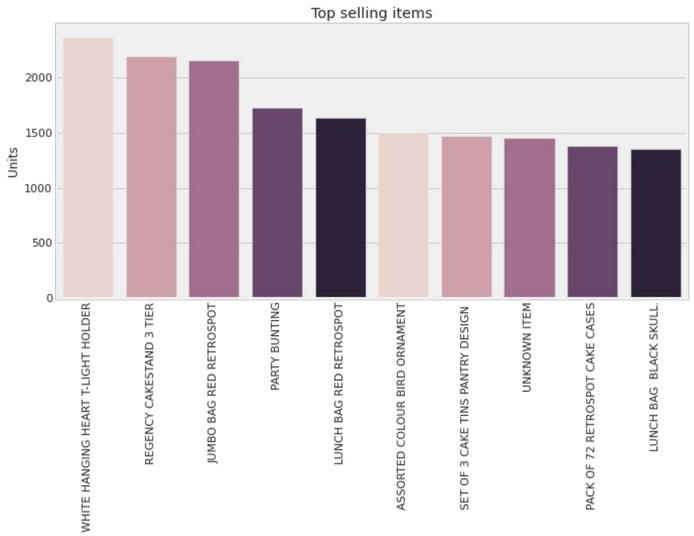
Methodology (data preparation - visualization - ML)

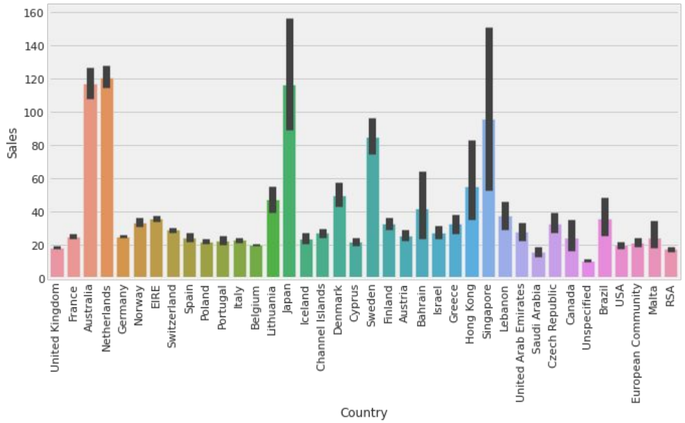
Data prep: Pandas, numpy

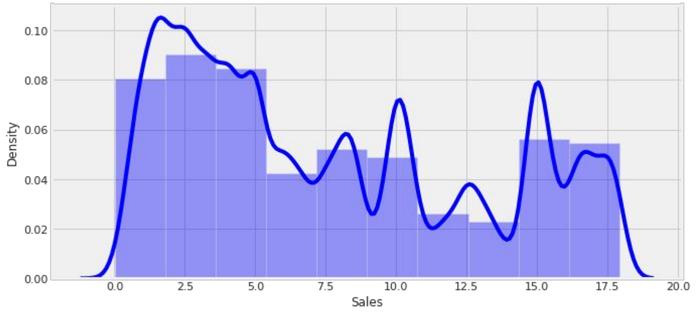
Visualisation: seaborn, matplotlib

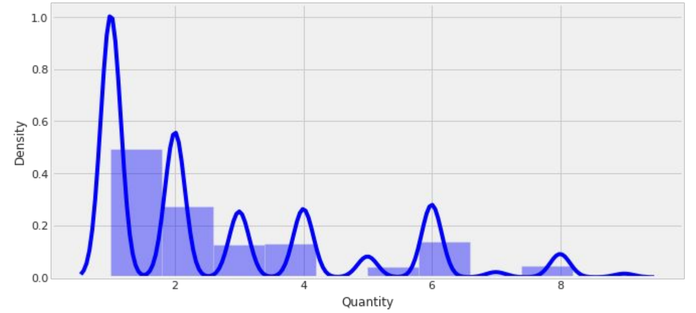
Forecasting: Linear regression, random forest

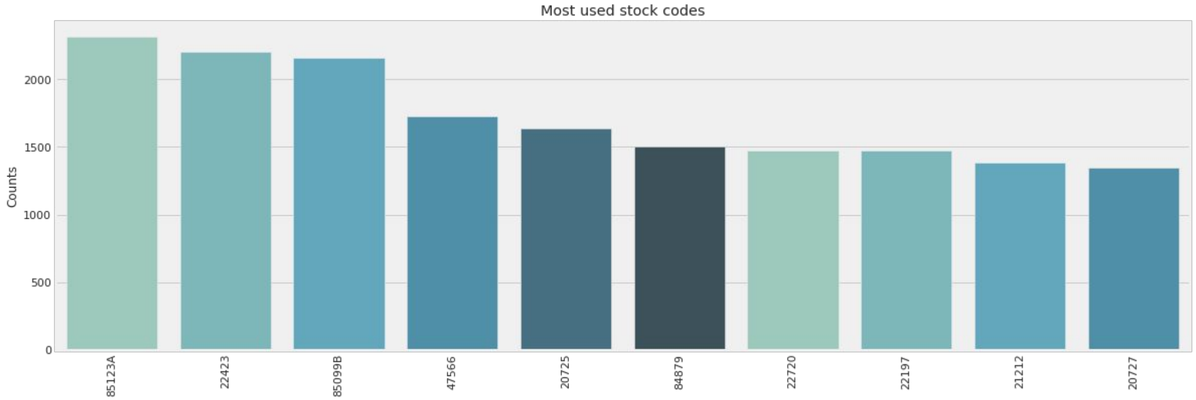
Dataset visualisation

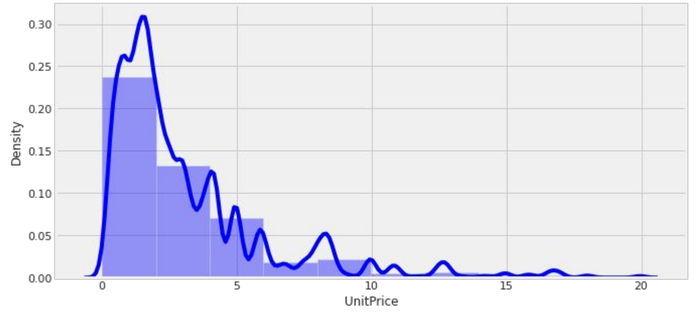












Sales forecasting: Linear regression and Random Forest output

Linear Regression: not precise with low r2 score

```
df_result = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
df_result
```

	Actual	Predicted
410279	8.29	38.295903
307883	16.50	9.196391
87541	2.50	1.321954
406268	5.90	7.065937
435894	30.00	17.772526
...	...	...
489106	4.68	10.853190
82441	3.29	6.989508
240406	34.80	24.537870
52477	19.80	20.846381
246991	23.00	22.608862

90218 rows × 2 columns

```
regressor.score(X_test, y_test)
```

```
0.4532204327432317
```

Random Forest: high r2 score implies high precision

```
df_result2 = pd.DataFrame({'Actual': y_test, 'Predicted': pred})
df_result2
```

	Actual	Predicted
410279	8.29	8.290000
307883	16.50	16.500000
87541	2.50	2.500000
406268	5.90	5.900000
435894	30.00	30.000000
...	...	...
489106	4.68	4.680000
82441	3.29	3.290002
240406	34.80	34.800000
52477	19.80	19.800000
246991	23.00	22.996100

90218 rows × 2 columns

```
[289] r2_score(pred,y_test)
```

```
0.979960619107674
```

ReDI School of Digital Integration

Thank you / Danke!

Your name / Ihre Name

COURSE / KURS