

Импорт библиотек

```
In [2]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from pylab import rcParams
rcParams['figure.figsize'] = 20, 10
```

Задача 1:

- Проведите аналитику полученных данных
- Подготовьте данные для регрессионной модели
- Обучите регрессионную модель и осуществите прогноз на последний месяц для поля SALES
- Оцените качество модели с помощью функции metric

Метрика для оценки Вашей модели

```
In [168]: def metric(y_true, y_pred):
return np.sum(np.abs(y_true - y_pred)) / np.sum(y_pred) * 100
```

Данные

```
In [2]: df = pd.read_csv(r'data_regression_for_task.csv')
```

```
In [3]: df.head(5)
```

Out[3]:

	YEAR	MONTH	CONTRAGENT	ARTICLE_CODE	ARTICLE_NAME	ARTICLE_GROUP	SALES	€
0	2017	4	ROYAL WINE CORP	100200	GAMLA CAB - 750ML	WINE	0.0	
1	2017	4	SANTA MARGHERITA USA INC	100749	SANTA MARGHERITA P/GRIG ALTO - 375ML	WINE	0.0	
2	2017	4	JIM BEAM BRANDS CO	10103	KNOB CREEK BOURBON 9YR - 100P - 375ML	LIQUOR	0.0	
3	2017	4	HEAVEN HILL DISTILLERIES INC	10120	J W DANT BOURBON 100P - 1.75L	LIQUOR	0.0	
4	2017	4	ROYAL WINE CORP	101664	RAMON CORDOVA RIOJA - 750ML	WINE	0.0	

Задача 2:

- Проведите аналитику полученных данных
- Подготовьте данные для модели классификации
- Обучите модель классификации и осуществите прогноз для отложенной выборки с учетом перекоса в данных TARGET
- Оцените качество модели с помощью функции `f1_score`

Данные

```
In [5]: df = pd.read_csv(r'data_classification_for_task.csv')
```

```
In [6]: df.head(5)
```

Out[6]:

	AGE	GENDER	FEATURE_1	FEATURE_2	FEATURE_3	FEATURE_4	FEATURE_5	FEATURE_6
0	63	1	3	145	233	1	0	150
1	37	1	2	130	250	0	1	187
2	41	0	1	130	204	0	0	172
3	56	1	1	120	236	0	1	178
4	57	0	0	120	354	0	1	163

