

# Тест

---

Привет! Это небольшой тест на проверку твоего знания Питона и базовое умение использовать Pandas для аналитики данных.

Удачи!

---

В исходном файле с данными «Test\_1.csv» представлена ежедневная информация по метрикам проекта в разбивке на регионы присутствия.

Комментарии по метрикам:

Регистрации\_organic– органические регистрации

Регистрации\_paid– регистрации, пришедшие по рекламе в интернете

Переходы\_paid – клики по рекламным сообщениям в интернете и переход в стор приложений

Orders\_all– количество заказов за день

Orders\_with\_coupon– количество заказов в день, совершенных с использованием промо-купонов

TVR –количество рейтингов на ТВ за день

Reach – количество людей, которые видели ТВ рекламу за день хотя бы 1 раз

OTS – количество контактов людей с рекламой на ТВ за день

Num\_of\_OOH – число сторон в кампании наружной рекламы

T – Средняя температура за день

RRR – число осадков за день

sss- число выпавшего снега

---

По ссылке <https://cloud.mail.ru/public/2Ekk/4Hsx3Ksp3> вам предоставлены данные по проекту «Эльбрус», представленному в регионах и информация о региональных ТВ-активностях

---

## Задание

---

*Вам необходимо оценить:*

- A) сколько платежей в каждом городе принесла ТВ-кампания (с обоснованием валидности расчетов и оценкой качества модели),
  - B) какие факторы имеют наибольший вес,
  - C) какие факторы вы бы добавили в модель, чего вам не хватает,
  - D) сделать кластеризацию городов по ТВ – эффекту. Сделайте обоснованные выводы в каких городах ТВ сработал лучше/хуже. Можно ли выделить группы городов, где наблюдаются схожие паттерны?
- 

[1]

```
import datetime

import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import plotly

import statsmodels
import statsmodels.tsa.arima_process
```

/usr/local/lib/python3.7/dist-packages/statsmodels/tools/\_testing.py:19: FutureWarning:

pandas.util.testing is deprecated. Use the functions in the public API at pandas.testing instead.

---

[2]

```
from google.colab import files

uploaded = files.upload()

for fn in uploaded.keys():

    print('User uploaded file "{name}" with length {length} bytes'.format(
        name=fn, length=len(uploaded[fn])))
```

---

[3]

```
data_for_model = pd.read_csv('Test_1.csv', sep=';', engine='python',
encoding='Windows 1251', decimal= ',')
```

---

[4]

```
data_for_model.head()
```

---

[10]

```
import seaborn as sns

sns.lineplot(data=data_for_model, x='Date', y='Orders_all')
```

---

[]

```
('01.01.2018', '31.12.2019')
```

---

---