

# **Деколонизация испанских музеев**

Илюхина Алина, Плишко Дарья (МЦМГН221)

# О проекте

- Курсовой проект о музеях Великобритании и Испании
- Мини проект для курса “Анализ данных” только о музеях Испании
- Что ещё можно сказать о собранных данных?
- Применение навыков анализа данных на куске дата сета большого проекта → внедрение в курсовой проект

# О данных

- **Откуда берём данные?** Сайты музеев
- **Каких музеев?**

## **Great Britain:**

British Museum  
Pitt Rivers Museum  
Cambridge Museum of Archaeology and Anthropology  
Bristol Museum & Art Gallery  
Bookery Gallerie

## **Spain:**

Museo de América  
Museo de Reina Sofia  
Museo del Prado  
Museo Arqueológico Nacional  
Museo Nacional de Antropología

- **Какие данные?** Информация о выставках, посвященных культуре бывших колоний Испании

# Таблица

Признаки: id, museum, type, date  
start/end, extension, year  
start/finish, region/2, country/2,  
name, subject, url, highlights,  
offline/online, admission fee,  
associated events

Что добавить? - Place, extended  
subjects

[https://docs.google.com/spreadsheets/  
d/1OHn2mbi8n6zprdGVnhq31flcGq\\_N  
2wPk4O\\_5jVeJLbs/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1OHn2mbi8n6zprdGVnhq31flcGq_N2wPk4O_5jVeJLbs/edit?usp=sharing)

# Гипотезы

## №1

H0 – объекты зависят от музеев, в которых они экспонируются.

H1 – объекты не зависят от музеев, в которых они выставляются.

## №2

H0 – объекты зависят от регионов, в которых они экспонируются.

H1 – объекты не зависят от регионов, в которых они выставляются.

## №3

H0 – регионы зависят от музеев, в которых они экспонируются.


H1 – регионы не зависят от музеев, в которых они выставляются.


# Методы

- Какие библиотеки использовали:

```
import numpy as np
import pandas as pd
import seaborn as sns
import scipy.stats as stats
from scipy.stats.contingency import association
```

- Проверка, как загрузились наши данные

 `df.info()`

 `<class 'pandas.core.frame.DataFrame'>`  
RangeIndex: 94 entries, 0 to 93  
Data columns (total 31 columns):

#	Column	Non-Null Count	Dtype
0	url	94 non-null	int64
1	country_of_museum	94 non-null	object

# Методы

- Что делали с категориальными данными

## Регионы

```
[ ] regions = list(set(df.region))
cat_regions = []
for i in df.region:
    n2 = regions.index(i) + 1
    cat_regions.append(n2)
df['cat_regions'] = cat_regions
```

```
[ ] regions
```

```
['Asia',
 'East Asia',
 'East Africa',
 'South Asia',
 'West Africa',
```

## Музеи

```
[ ] museums = list(set(df.Museum))
cat_museum = []
for i in df.Museum:
    n = museums.index(i) + 1
    cat_museum.append(n)
df['cat_museum'] = cat_museum
```

```
[ ] museums
```

```
['BM&AG', 'PR', 'MDA', 'MAA', 'MDP', 'BG', 'MNA', 'BM', 'MRS']
```

## Объекты

```
[ ] objects = list(set(df.objects_type))
cat_objects = []
for i in df.objects_type:
    n1 = objects.index(i) + 1
    cat_objects.append(n1)
df['cat_objects'] = cat_objects
```

```
[ ] objects
```

```
[nan,
 'artefacts',
 'armour',
 'archaeological pieces',
```

# Проверка

Для проверки наших гипотез мы используем критерий хи-квадрат  
Определяем нулевую и альтернативную гипотезы

[ ] #H0 - объекты зависят от музеев, в которых они выставляются  
#H1 - объекты не зависят от музеев, в которых они выставляются

[ ] #H0 - объекты зависят от регионов, в которых они выставляются  
#H1 - объекты не зависят от регионов, в которых они выставляются

[ ] #H0 - регионы зависят от музеев, в которых выставляются  
#H1 - регионы не зависят от музеев, в которых они выставляются



# Проверка гипотез

Гипотеза №1

```
0.08046075176878459 < 0.05  
#pvalue > 50%  
# объекты не зависят от музеев – типы объектов могут быть равномерно распределены по музеям  
False
```

Объекты не зависят от музеев в  
которых они выставляются.

Они могут быть равномерно  
распределены по музеям.

# Проверка гипотез

## Гипотеза №2

```
5.636960409641096e-07 < 0.05  
# pvalue < 50%  
# объекты зависят в от регионов, которые выставляются на выставках  
True
```

Объекты зависят от регионов в которых выставляются.

- корреляция между объектами и регионами сильнее, чем корреляция между регионами и музеями (при рассмотрении результатов значения pvalue можно заметить, что при проверке 2 гипотезы он больше чем при проверке 3)

```
1.6881669537033782e-11 < 5.636960409641089e-07  
True
```

# Проверка гипотез

Гипотеза №3

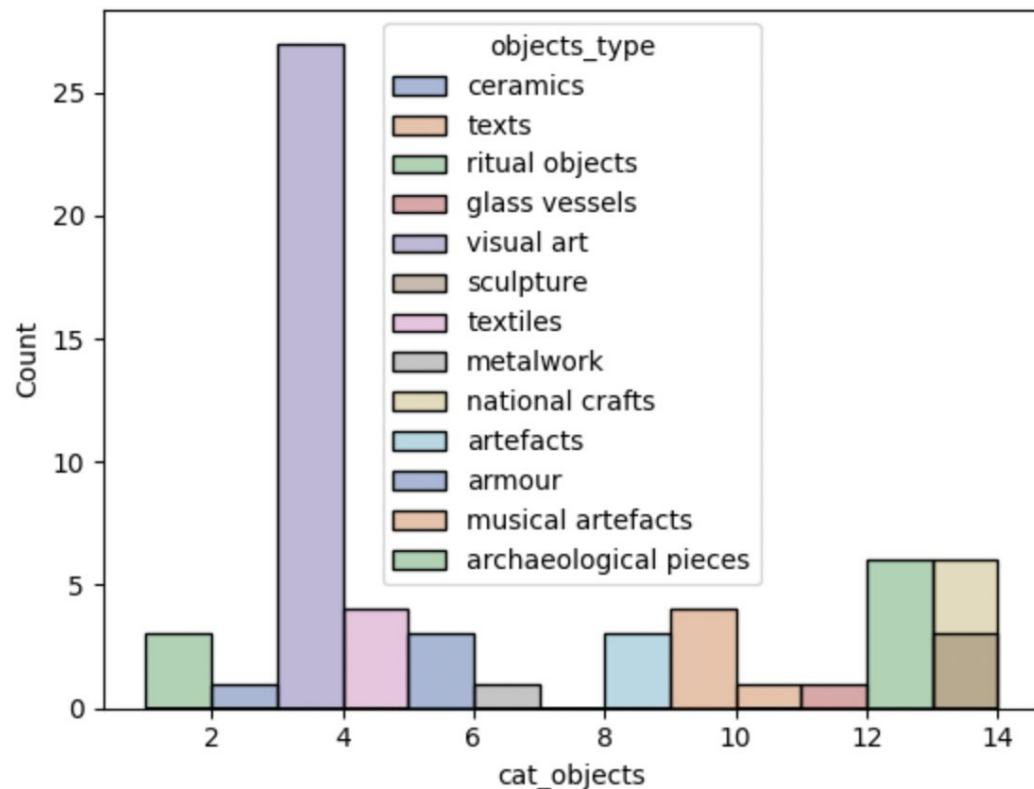
```
▶ 1.6881669537033782e-11 < 0.05  
# pvalue < 50%  
# регионы зависят от музеев, в которых выставляются  
True
```

Регионы зависят от музеев, в которых они выставляются.

(Например, музеи Испании чаще выставляют культуру Латинской Америки, а музеи Великобритании культуру Африки)

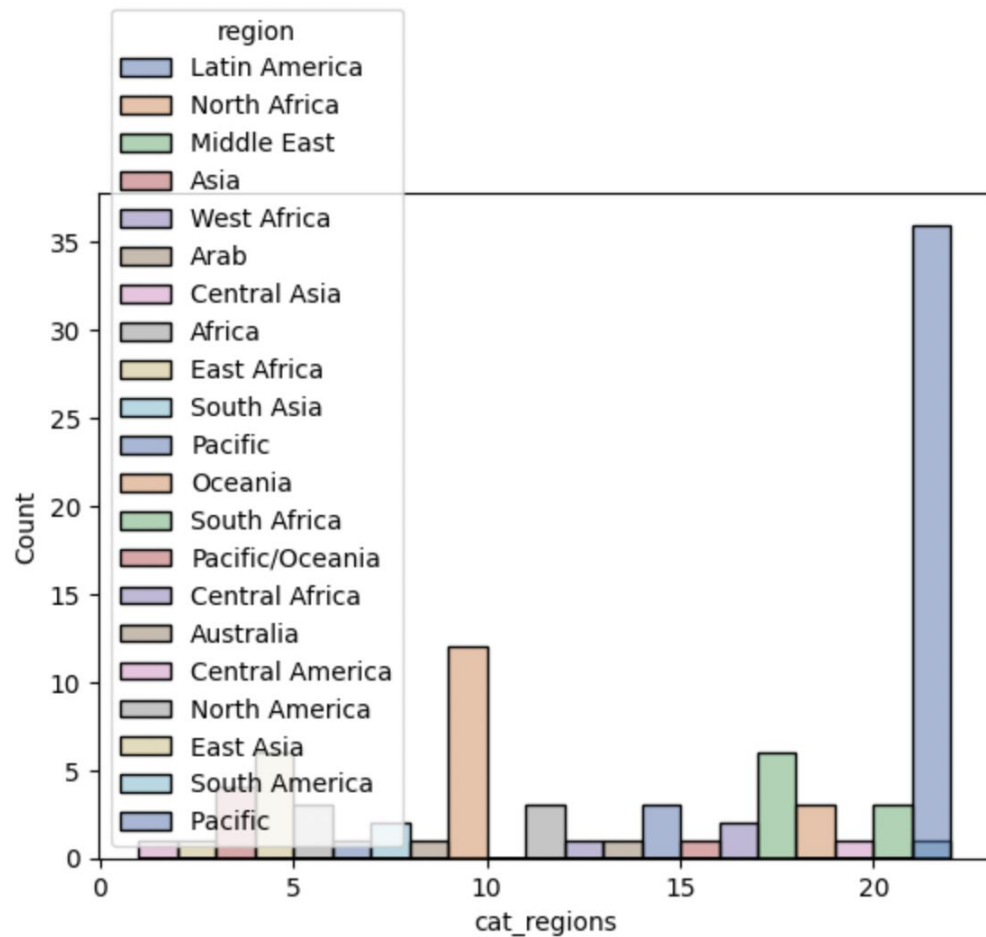
# Что еще выявили

Популярные и не очень объекты



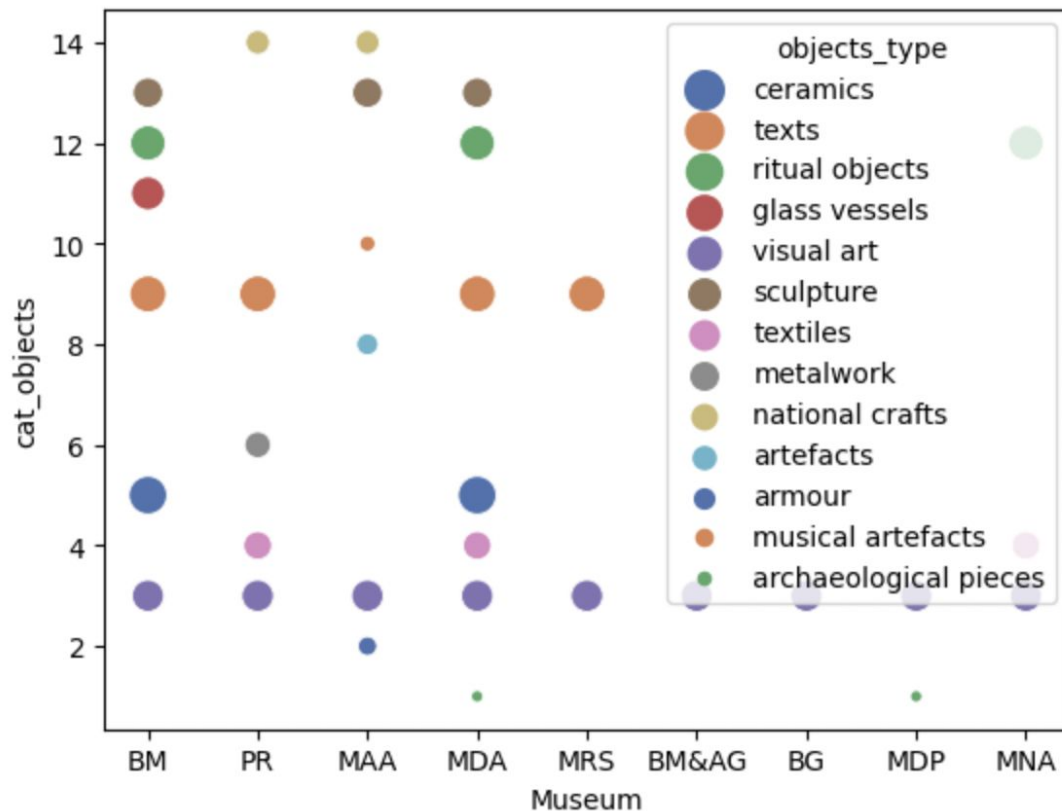
# Что еще выявили

Популярные и не очень регионы



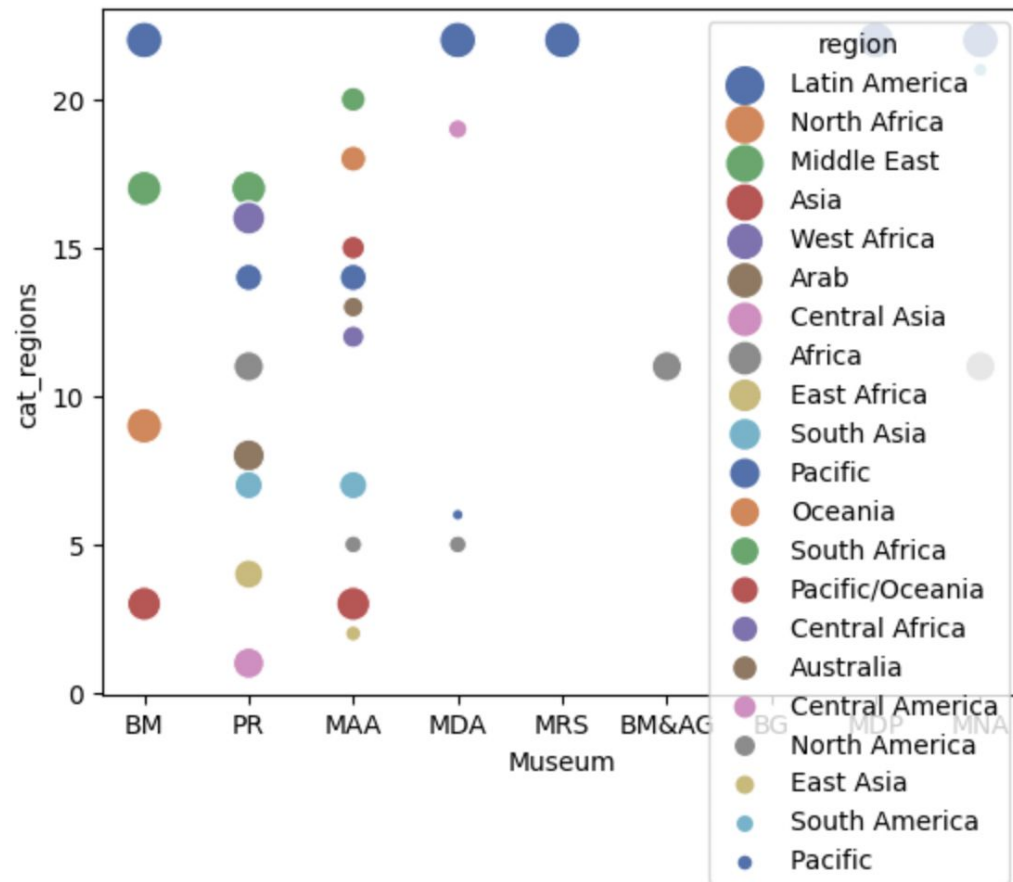
# Что еще выявили

Корреляция музеи-объекты



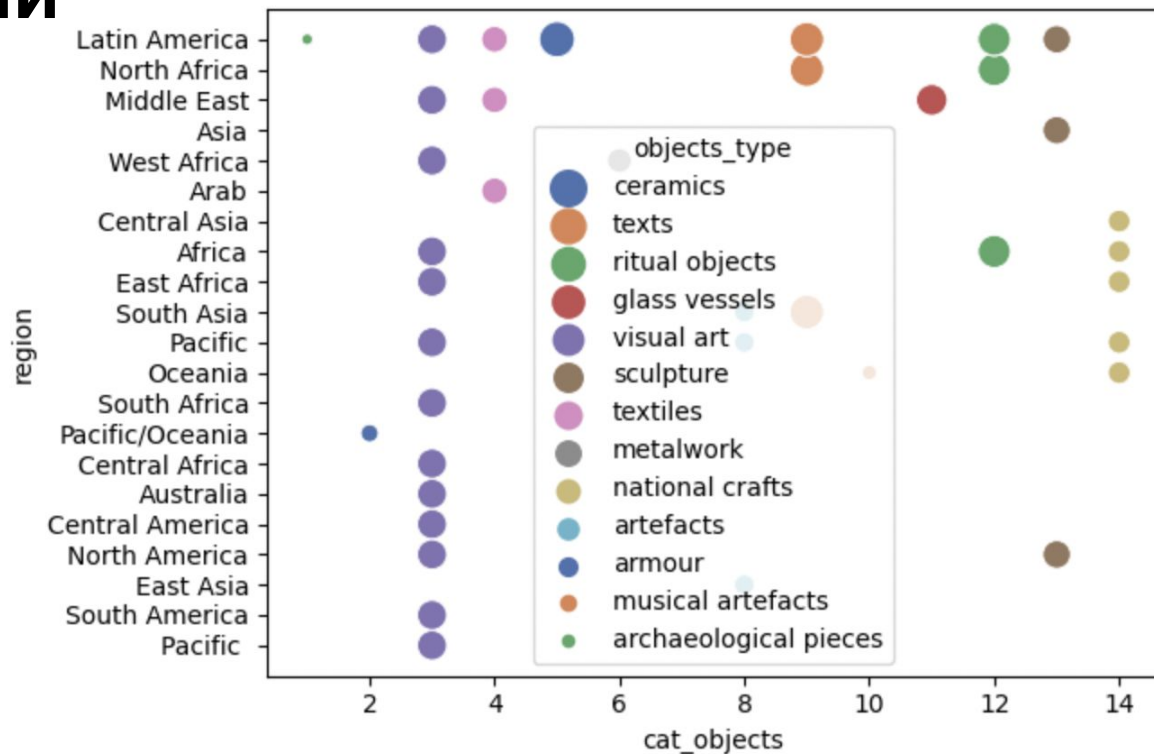
# Что еще выявили

Корреляция музей-регионы



# Что еще выявили

Корреляция регионы-объекты





**Презентация окончена!**

**Спасибо за внимание**