

---

---

# 20. Анализ и Youtube Faces

— 5 януари 2023 —

---

---

# За какво си говорихме предния път

- Основни разлики между Data Science, AI и ML
- Разни полезни инструменти (но не заместители)
- Доста за NumPy
- Доста и за Pandas
- Започнахме да разглеждаме данни

# Малко въпроси

```
a = np.array([0,1,2,3,4,5,6,7,8,9])
```

```
b = a.reshape(3,3)
```

```
b
```

```
ValueError: cannot reshape array of size 10 into shape (3,3)
```

# Малко въпроси

```
a = np.arange(20).reshape(4, 5)
```

```
a[1:2, 3]
```

---

```
array([[ 0,  1,  2,  3,  4],  
       [ 5,  6,  7,  8,  9],  
       [10, 11, 12, 13, 14],  
       [15, 16, 17, 18, 19]])
```

# Малко въпроси

```
arr1 = [0, 1, 2, 3, 4, 5]
```

```
arr2 = [6, 7, 8, 9, 10, 11, 12]
```

```
s1 = pd.Series(arr1)
```

```
s2 = pd.Series(arr2)
```

```
s2.add(s1)
```

```
0    6.0  
1    8.0  
2   10.0  
3   12.0  
4   14.0  
5   16.0  
6     NaN  
dtype: float64
```

# Процес

- Някаква цел- Какво ще се опитваме да предсказваме? Да обясним и тн?
- Данни- Дефиниране и обработка
- Моделиране- Според зависи това може да е нещо малко и бързо, а може и да е нещо, върху което да работим месеци

# Цел за деня

- Ще се върнем на данните от предния път
- Ще направим модел, който да разпознава **накъде гледат** лица

# Почистване и обработка на налични данни

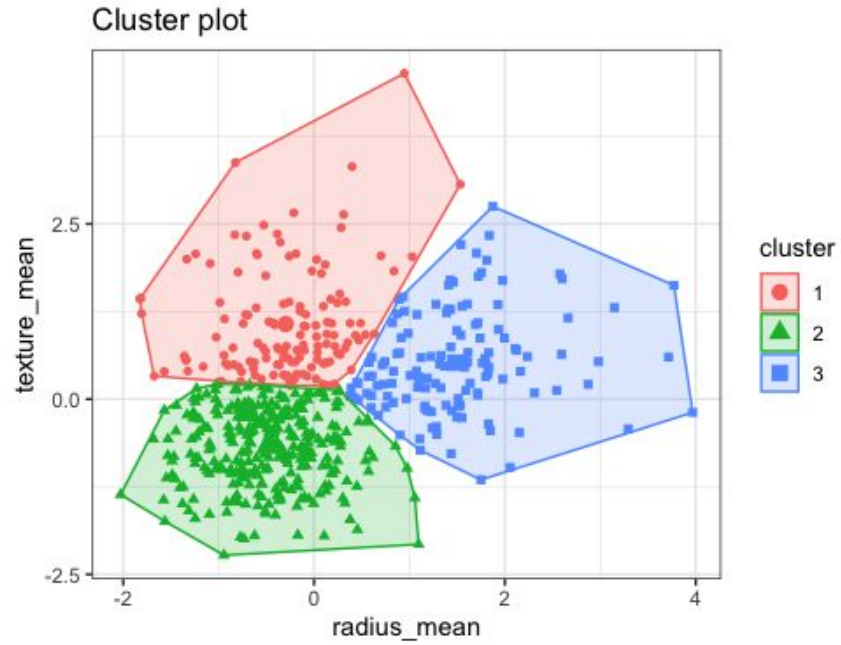
- Outliers- според зависи от ситуацията се махат
- Трансформации
- Корелация
- Липсващи данни
- Създаване на нови



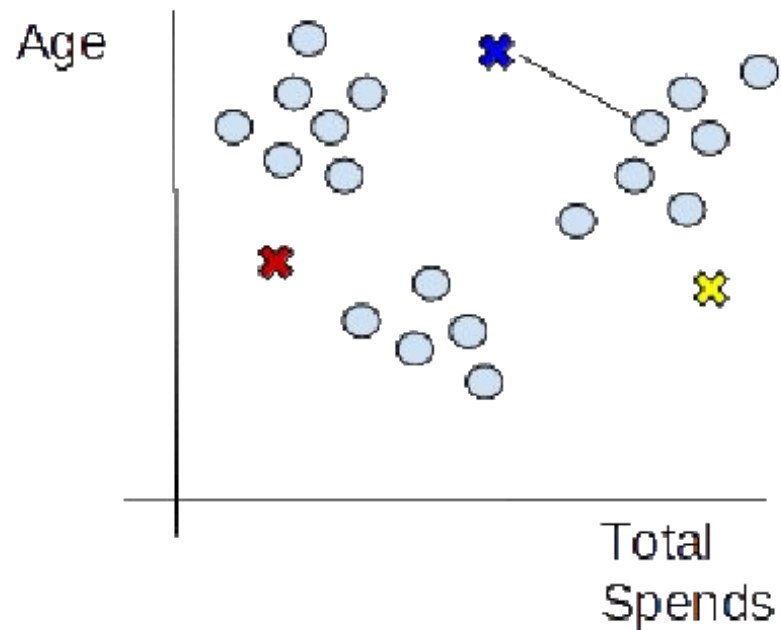
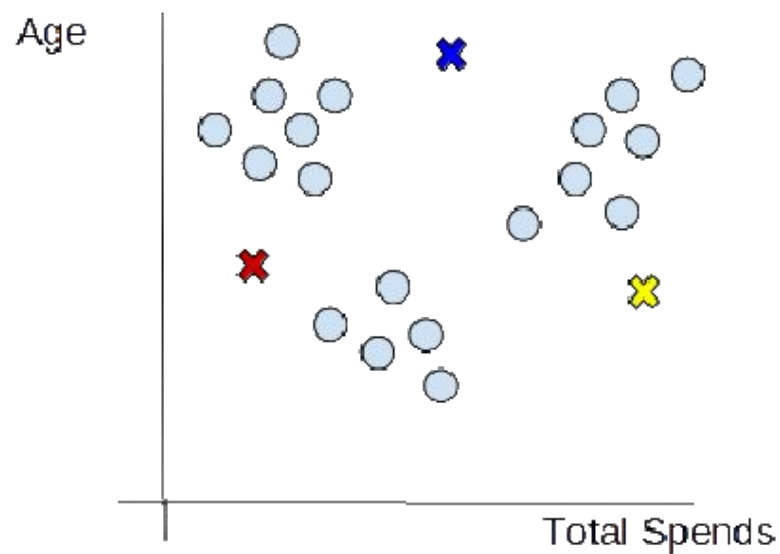
# YouTube Faces

- Какво представляват данните?- фреймове от ютуб видеа с предварително зададени точки на основните черти на лица
- Крайна цел?- да ги класифицираме, спрямо тяхната форма
- Ще използваме K-means.

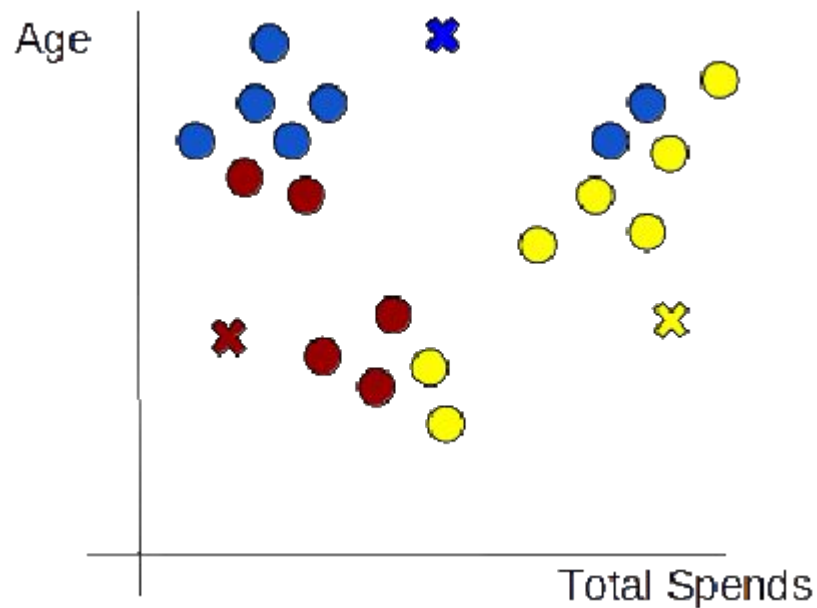
# K-means



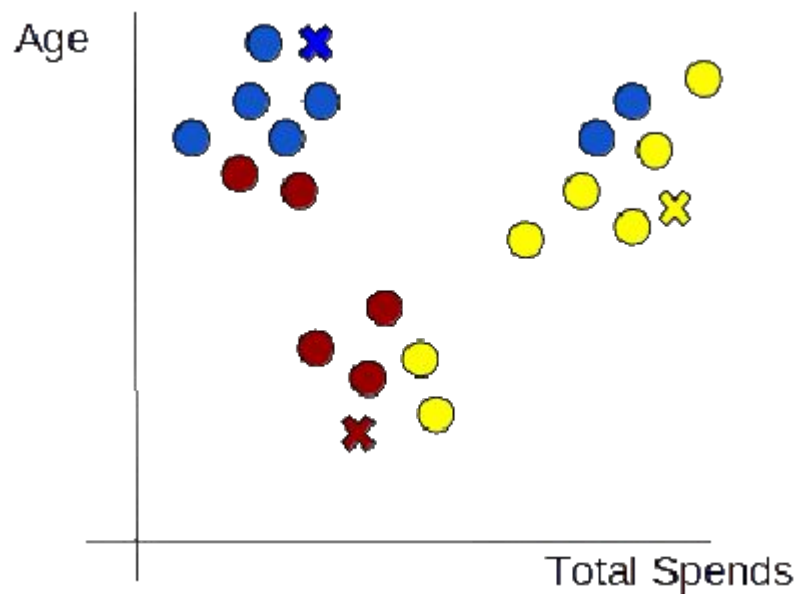
## K-means- как работи?



# K-means- как работи?



# K-means- как работи?





# Линкове към кода и (част от) данните

<https://drive.google.com/drive/folders/1QHLV7BOyqPULMJNYQYwMmoKBdPj2h0SW>

**Въпроси?**