

Университет ИТМО

Практическая работа №2
по дисциплине «Визуализация и моделирование»

Автор: Винникова Ева Михайловна

Поток: ВИМ 1.2

Группа: К3222

Факультет: ИКТ

Преподаватель: Чернышева А.В.

Санкт-Петербург, 2021 г.

Датасет содержит информацию о женских корейских поп группах.

Столбец	Данные	Тип	Шкала
Name	название группы	строка	Номинальная
Gender	сокращенное название группы	строка	Номинальная
Korean name	название группы на корейском	строка	Номинальная
Debut	дата дебюта	дата	Относительная
Company	компания группы	строка	Номинальная
Members	количество участников в группе	число	Относительная
Orig.membr.	планируемое количество участников	число	Относительная
Fanclub name	название фанатской базы	строка	Номинальная
Active	активна ли группа на данный момент	строка	Номинальная

1. Какой процент групп имеют собственное название фандома?

Вывод: Только 35 процентов имеют название фандома, это говорит о том, что практика создания имени все еще не так популярна

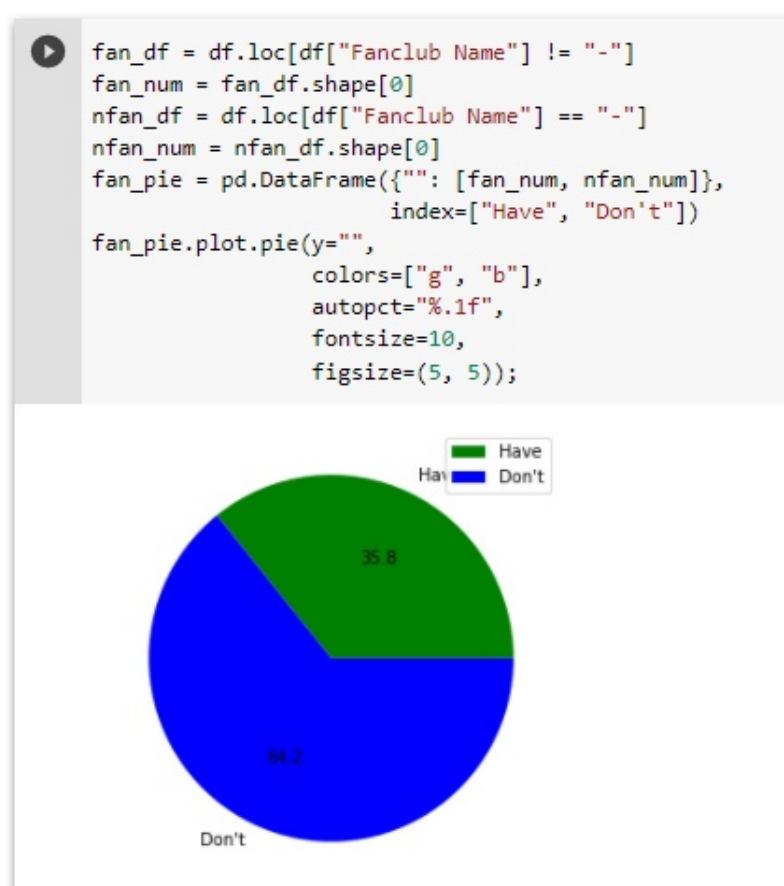


Рисунок 1 - Наличие названия фандома

2. Какой процент групп больше не активны?

Вывод: Из всех групп только 31 процент неактивны, это значит, что количество активных групп только увеличивается.

```

act_df = df.loc[df["Active"] == "Yes"]
act_num = act_df.shape[0]
nact_df = df.loc[df["Active"] == "No"]
nact_num = nact_df.shape[0]
act_pie = pd.DataFrame({"": [act_num, nact_num]},
                        index=["Yes", "No"])
act_pie.plot.pie(y="",
                 colors=["c", "m"],
                 autopct="%.1f",
                 fontsize=10,
                 figsize=(5, 5));

```

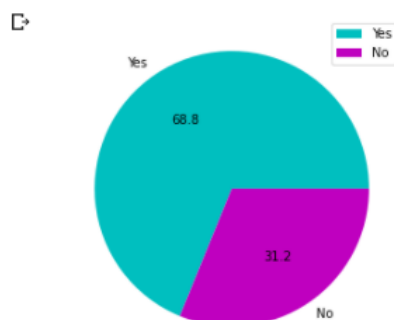


Рисунок 2 - Активность групп

3. Каких групп больше женских или мужских?

Вывод: количество женских и мужских групп примерно одинаково.

```

male_df = df.loc[df["Gender"] == "M"]
m_num = male_df.shape[0]
fem_df = df.loc[df["Gender"] == "F"]
f_num = fem_df.shape[0]
age_pie = pd.DataFrame({"": [m_num, f_num]},
                        index=["M", "F"])
age_pie.plot.pie(y="",
                 colors=["c", "m"],
                 autopct="%.1f",
                 fontsize=10,
                 figsize=(5, 5));

```

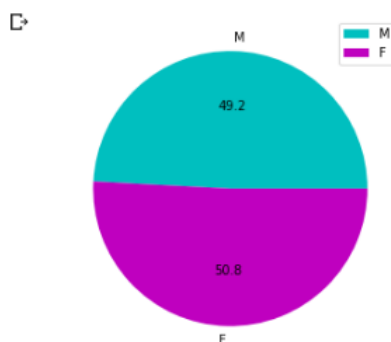


Рисунок 3 - Распределение пола

4. Какие компании создали наибольшее количество групп?

Вывод: SM создала больше всего групп, на втором месте JYP и т. д. На графике можно увидеть распределение групп по компаниям.

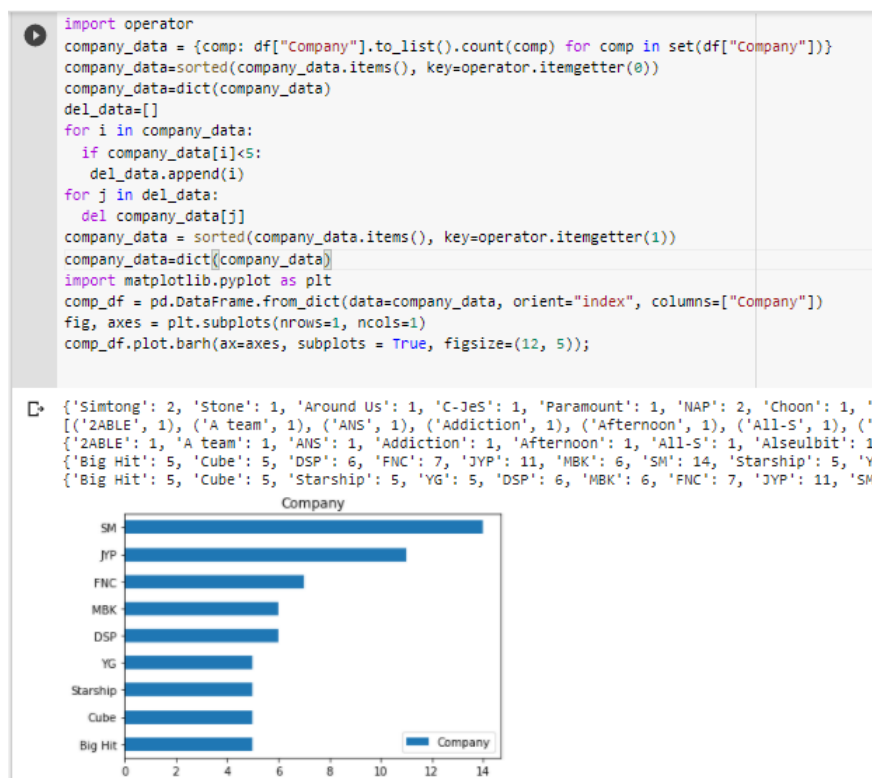


Рисунок 4 - Распределение по компаниям

5. Соотношение женских и мужских групп в компании SM

Вывод: В компании SM женских групп больше, чем мужских.



Рисунок 5 - Распределение по полу в SM

6. Какое самое популярное число участников в к-поп группе?

Вывод: Самое частое число участников 5, распределение по остальным числам видно на графике.

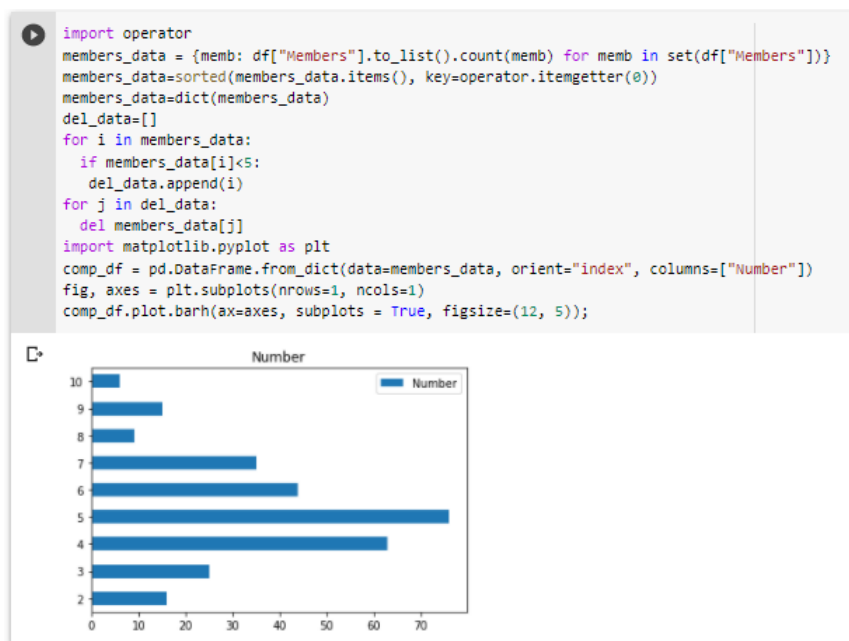


Рисунок 6 - Число участников

7. Как изменялось количество выпущенных групп с годами?

Вывод: Количество относительно планово возрастало с каждым годом, но в 2019 и 2020 резко уменьшилось из-за коронавируса.

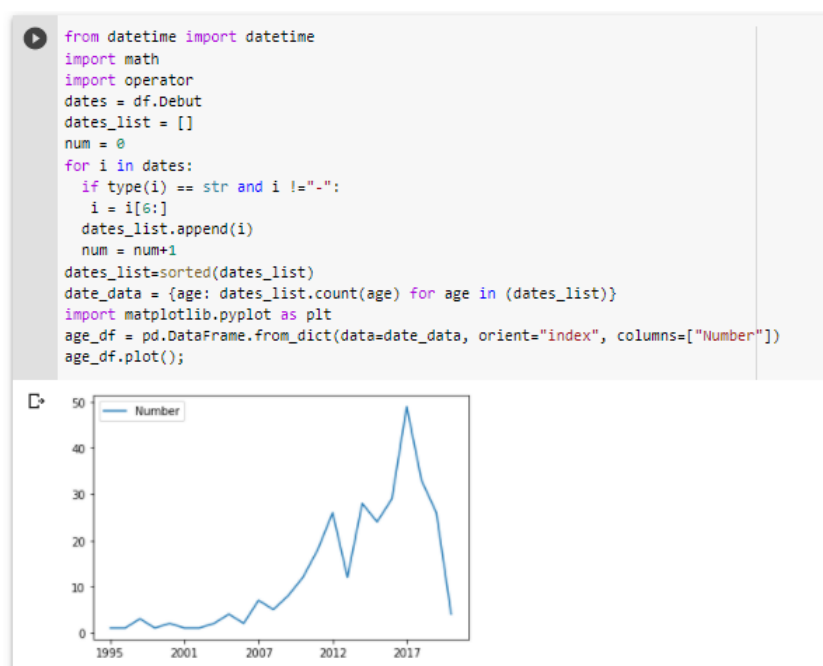


Рисунок 7 - Количество новых групп

8. Как изменялось количество распущенных групп с годами?

Вывод: Самый пик распадов был в 2017 и 2019 годах.

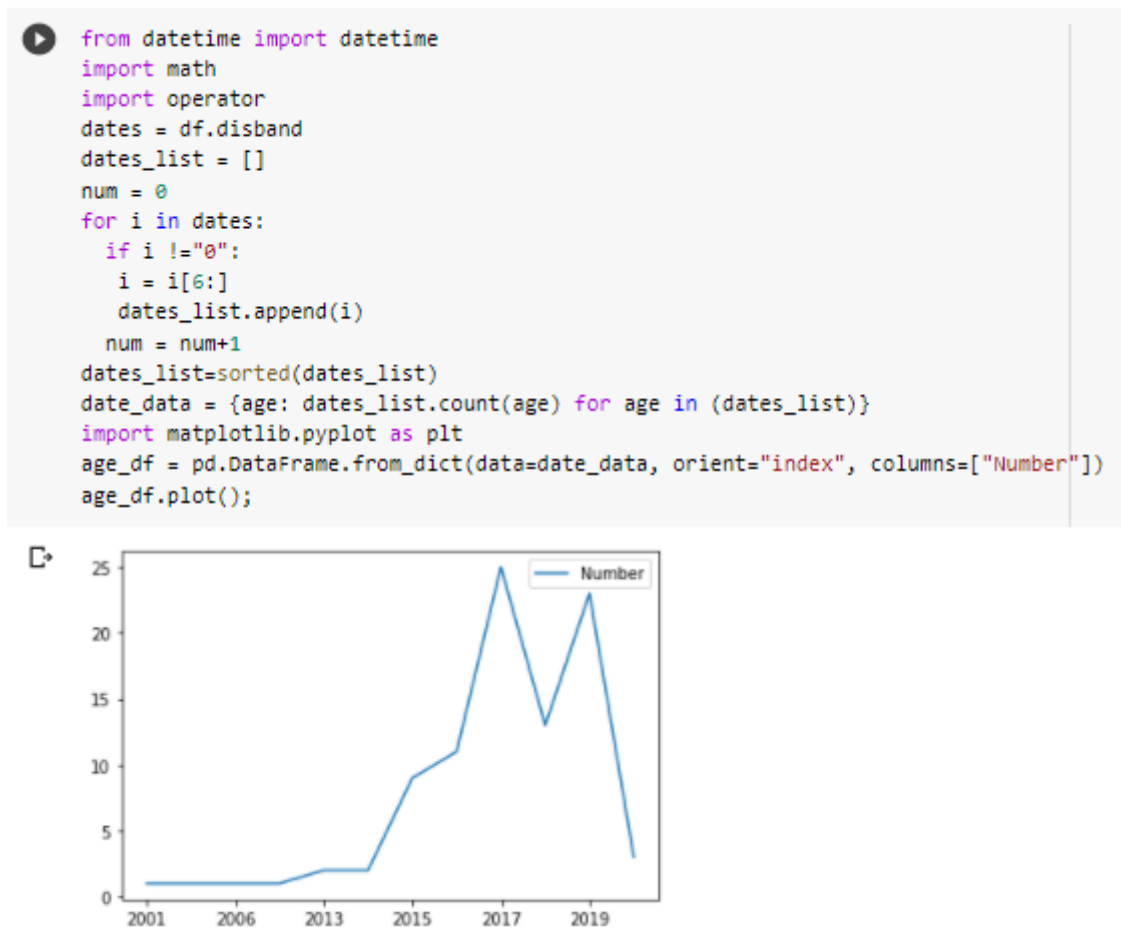


Рисунок 8 - Количество распавшихся групп

9. Какой процент групп у которых изменилось количество участников по сравнению с изначальным?

Вывод: только у 35 процентов изменилось количество участников, это означает что оно изменяется не очень часто.

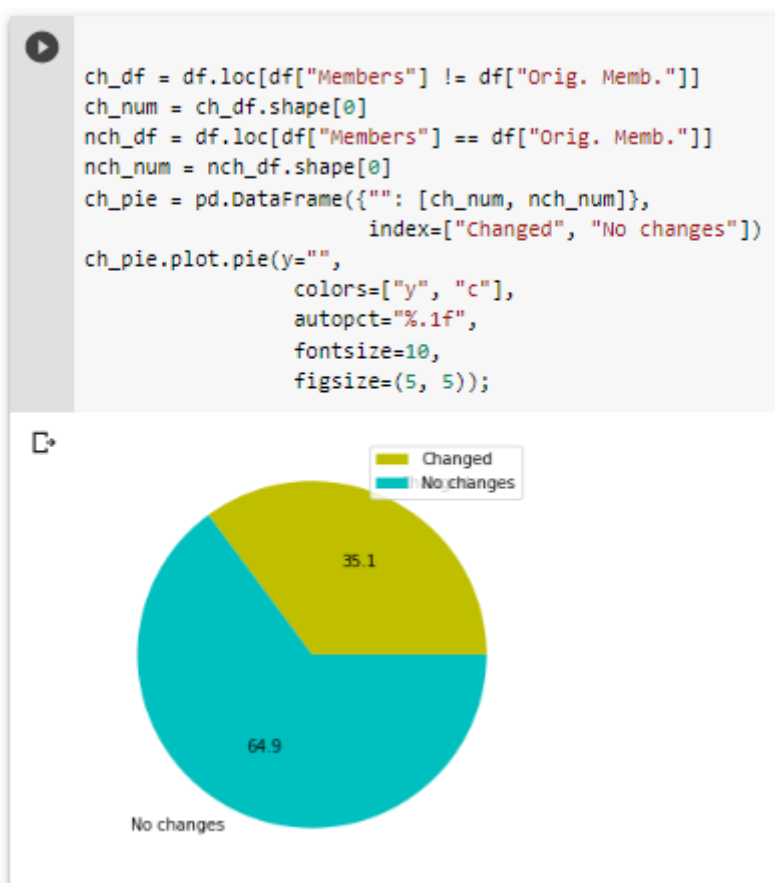


Рисунок 9 - Изменение количества участников

10. Сколько живут к-поп группы?

Вывод: Группы обычно живут 3-4 года.

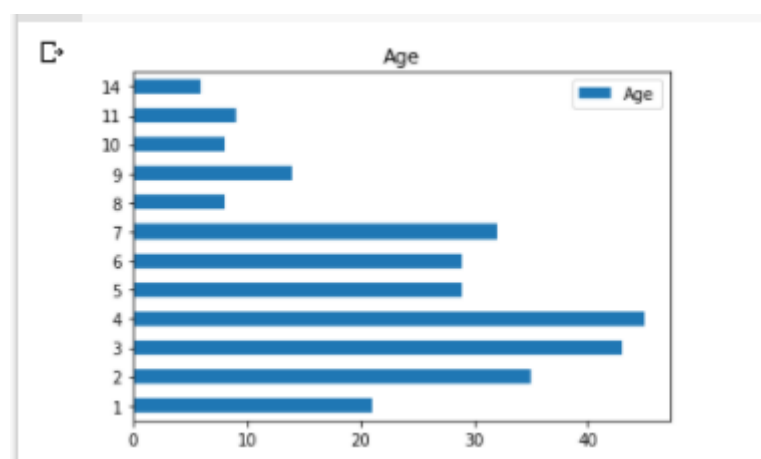


Рисунок 10 - Жизнь групп