# Семинар 8. Построение графиков и визуализация (деловая игра)

## Организация игры

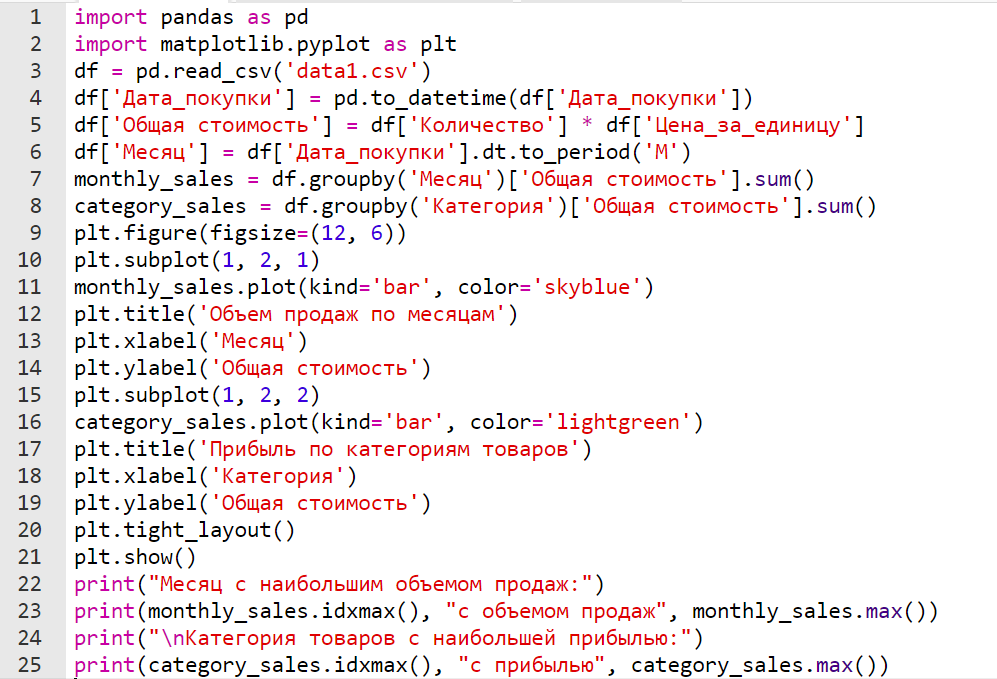
**Разделитесь на команды по 3-5 человек. Работа выполняется командой.** В данной работе вам будет нужно проанализировать возможности Python по обработке конкретной информации, научиться использовать библиотеки pandas и matplotlib для анализа и визуализации данных, а также развивать навыки принятия решений в бизнесе.

Запишите участников группы:

|  |
| --- |
| Москалев Юрий  Борисовская Алина  Галкина Полина  Головлёва Дарья |

В данном задании вам нужно разобраться с одним готовым примером обработки данных, а потом самостоятельно обработать предложенные данные.

Задание 1. Данные для задачи находятся в файле data1.csv. Загрузите файл в Trinket (https://trinket.io/embed/python3). Ниже представлен готовый код на Python анализа данных в двух вариантах (скриншот с номерами строк и в виде текста, для быстрой вставки в Trinket):



|  |
| --- |
| import pandas as pd  import matplotlib.pyplot as plt  df = pd.read\_csv('data1.csv')  df['Дата\_покупки'] = pd.to\_datetime(df['Дата\_покупки'])  df['Общая стоимость'] = df['Количество'] \* df['Цена\_за\_единицу']  df['Месяц'] = df['Дата\_покупки'].dt.to\_period('M')  monthly\_sales = df.groupby('Месяц')['Общая стоимость'].sum()  category\_sales = df.groupby('Категория')['Общая стоимость'].sum()  plt.figure(figsize=(12, 6))  plt.subplot(1, 2, 1)  monthly\_sales.plot(kind='bar', color='skyblue')  plt.title('Объем продаж по месяцам')  plt.xlabel('Месяц')  plt.ylabel('Общая стоимость')  plt.subplot(1, 2, 2)  category\_sales.plot(kind='bar', color='lightgreen')  plt.title('Прибыль по категориям товаров')  plt.xlabel('Категория')  plt.ylabel('Общая стоимость')  plt.tight\_layout()  plt.show()  print("Месяц с наибольшим объемом продаж:")  print(monthly\_sales.idxmax(), "с объемом продаж", monthly\_sales.max())  print("\nКатегория товаров с наибольшей прибылью:")  print(category\_sales.idxmax(), "с прибылью", category\_sales.max()) |

Объясните назначение кода, который находится в строках с номерами:

|  |  |
| --- | --- |
| **Номер строки** | **Назначение кода** |
| 3 | Чтение файла |
| 4 | Преобразование столбца «Дата\_покупки» в формат базы данных для упрощения работы |
| 5 | Задание формулы для расчёта общей стоимости |
| 6 | Создание нового столбца «Месяц» на основе даты покупки |
| 7 | Рассчитывает общий объём продаж по месяцам |
| 8 | Рассчитывает суммарную прибыль по категориям товаров |
| 9-14 | Создание столбчатой диаграммы, размещение его в первой из двух ячеек и установление размеров и установка заголовка, легенды и цвета столбиков |
| 15-20 | Создание второй столбчтой диаграммы и размещение уже во второй ячейке |
| 23 | Определение месяца с наибольшими продажами и её вывод |
| 25 | Определение категории товаров с наибольшей прибылью и её вывод |

Задание 2. В этом задании вам нужно использовать pandas для анализа данных. Вы можете использовать пример, разобранный выше в качестве подсказок. Загрузите данные из файла data2.csv. Используя pandas найдите ответы на следующие вопросы:

2.1. Какой товар продается лучше всего? (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

2.2. Какой регион приносит наибольшую прибыль? (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

2.3. Какой канал продаж наиболее эффективен? (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

2.4. Выведите информацию по регионам в порядке убывания **общих сумм по продажам** в каждой категории. (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

Задание 3. В этом задании вам нужно использовать matplotlib для визуализации данных. Продолжите работать с файлом data2.csv. С помощью matplotlib визуализируйте следующую информацию:

3.1. График продаж по регионам. (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

3.2. Гистограмма продаж по товарам. (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

3.3. Круговая диаграмма, показывающая долю каждого канала продаж. (вставьте скриншот с кодом и результатом работы):

|  |
| --- |
|  |

Задание 4. На основе проведенных ранее анализа и визуализаций, предложите стратегию для увеличения продаж. Это может включать (только в качестве примера):

* Изменение цен на определенные товары.
* Увеличение рекламы в регионах с низкими продажами.
* Расширение онлайн-продаж.
* И т.д.

|  |
| --- |
| Я считаю, целесообразным для увеличения продаж будет является: развитие филиала в сфере рекламы, понижение цен на технику, расширение ассортимента в онлайн-магазине компании и разработка новой функций, например, доставки. |

Задание 5. Составьте короткую презентацию по проделанной работе. Презентация должна включать:

* Основные выводы из анализа.
* Визуализации, поддерживающие их выводы.
* Предложенные меры для увеличения продаж.

Загрузите презентацию вместе с файлом отчета.