**Отчёт**

**Анализ данных (*EDA*)**

**Цель анализа:** Исследовать предоставленный датасет с информацией о продажах для выявления основных закономерностей, распределений и взаимосвязей.

**Этапы анализа:**

**1. Загрузка и первичный осмотр данных**

* Были загружены данные, в которых содержатся следующие колонки: *Order ID, Product, Quantity Ordered, Price Each, Order Date, и Purchase Address*.
* Выявлены пропуски, которые были удалены.

**2. Обработка пропущенных значений**

* Пропуски встречались в столбцах *Quantity Ordered и Price Each*.
* Эти строки были удалены, чтобы избежать искажений в последующем анализе.

**3. Анализ распределения переменных**

* Построена гистограмма для столбца Price Each (цена за единицу товара).

**4. Исследование корреляций**

* Была исследована корреляция между Price Each и Quantity Ordered.
* Коэффициент корреляции оказался низким, что говорит об отсутствии сильной линейной зависимости между этими переменными.

**5. Выявление выбросов**

* Построен boxplot для столбца Quantity Ordered, чтобы выявить аномалии в количестве заказов.
* Были замечены выбросы, указывающие на редкие случаи аномально больших заказов.

**6. Изучение категориальных переменных**

* Исследована популярность продуктов с помощью барчарта, показывающего частоту продаж различных товаров.
* Определены самые популярные и непопулярные товары.

**7. Визуализация результатов**

* На итоговой визуализации представлены:
  + Гистограмма распределения цен.
  + Диаграмма рассеяния количества и цены.
  + Ящик с усами для анализа выбросов.
  + Диаграмма частоты продаж продуктов