МЕЖГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ «БЕЛОРУССКО-РОССИЙСКИЙ УНИВЕРСИТЕТ»

Кафедра “Программное обеспечение информационных технологий”

Отчет

По лабораторной работе №2

БАЗЫ ЗНАНИЙ И ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ В СИСТЕМАХ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

Выполнили:

Cтуденты группы ИСиТ-221

Мельников А.В, Астапович A.А \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

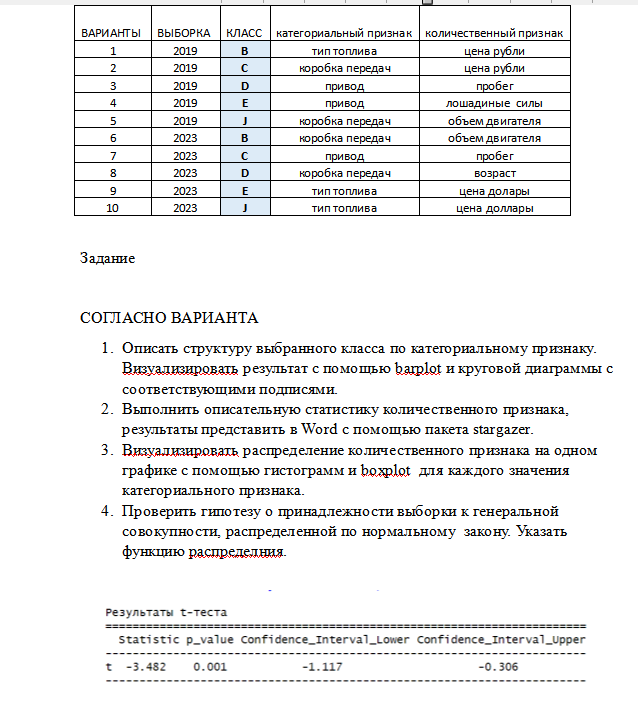
Проверил преподаватель

Ливинская В.А.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Могилев 2025

***Цель***: провести простой анализ данных, состоящий из описательной статистики и визуализации. освоить основные распространенные типы графиков в R, приобрести основные навыки работы с пакетом ggplot2 в R

Вариант 1

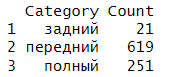


**1)** Описать структуру выбранного класса по категориальному признаку. Визуализировать результат c помощью barplot и круговой диаграммы с соответствующими подписями.

Выборка по категориальному признаку и классу авто– привод, класс - С.

Количество экземпляров в выборке – 891.

Распределение приводов в выборке:



Гистограмма:



Круговая диаграмма:



**2)** Выполнить описательную статистику количественного признака, результаты представить в Word с помощью пакета stargazer.

Визуализировать распределение количественного признака на одном графике с помощью гистограмм и boxplot для каждого значения категориального признака.

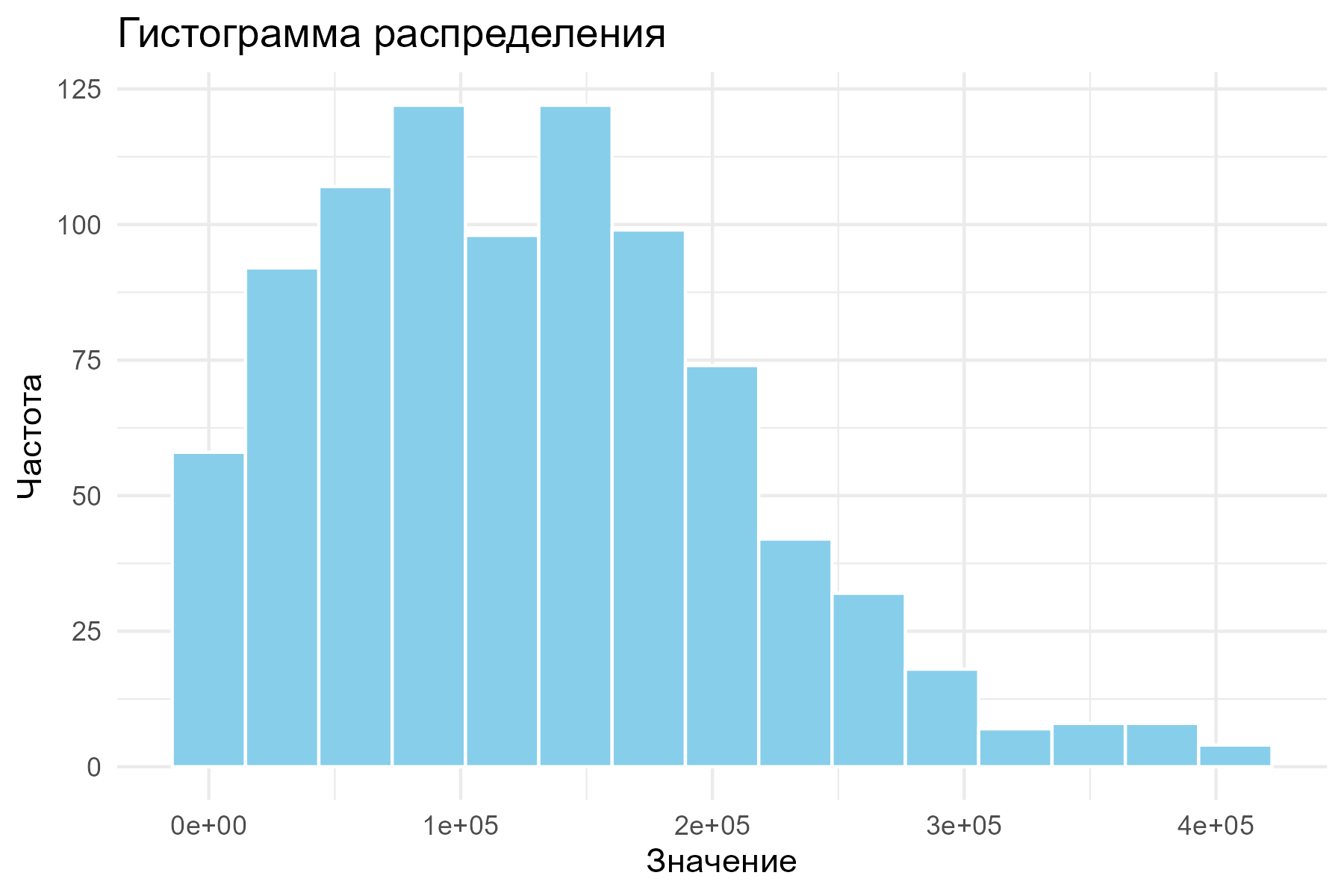
Данные выведенные в Word:

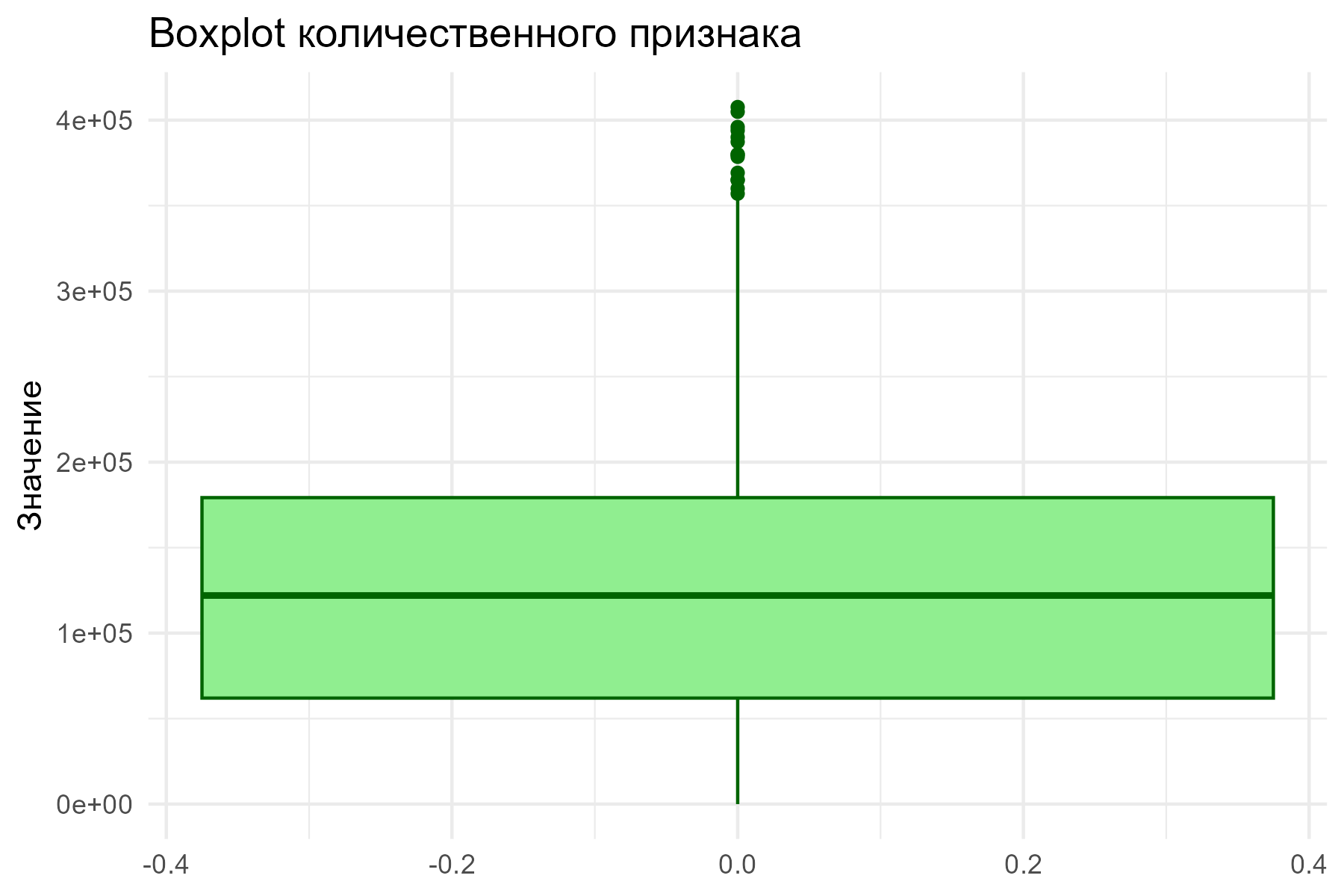
1. Анализ количественного признака

Описательная статистика количественного признака

| **Показатель** | **Значение** |
| --- | --- |
| Наблюдения | 891.00 |
| Минимум | 1.00 |
| Максимум | 407,664.00 |
| Среднее | 128,062.85 |
| Медиана | 122,000.00 |
| Станд. отклонение | 83,290.35 |
| Дисперсия | 6,937,282,350.64 |

## Визуализации



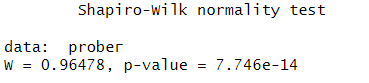


**3)** Проверить гипотезу о принадлежности выборки к генеральной совокупности, распределенной по нормальному закону. Указать функцию распределния.

Вывод описательной статистики:

****

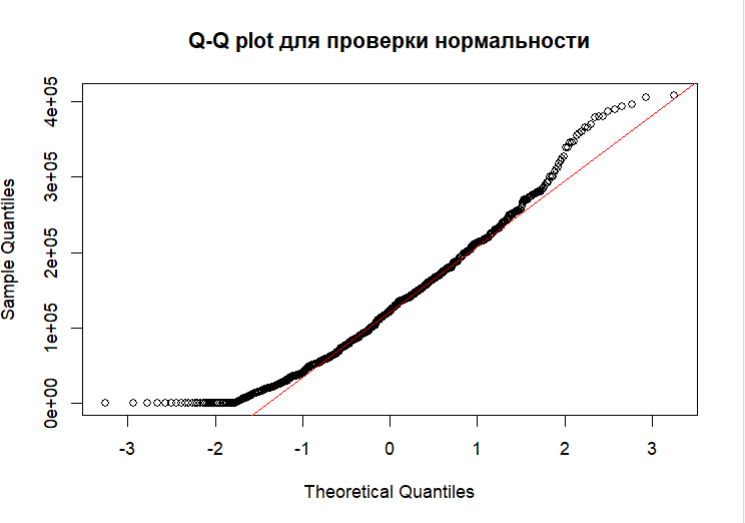
Проверку гипотезы провели с помощью теста Шапиро-Уилка на нормальность:



Гистограмма распределения пробега:

****

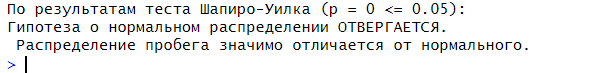
Q-Q plot для проверки на нормальность:

****

Вывод:

Поскольку тест Шапиро-Уилка выдает значение p<0.5,

следовательно гипотеза о нормальном распределении отвергается.

****