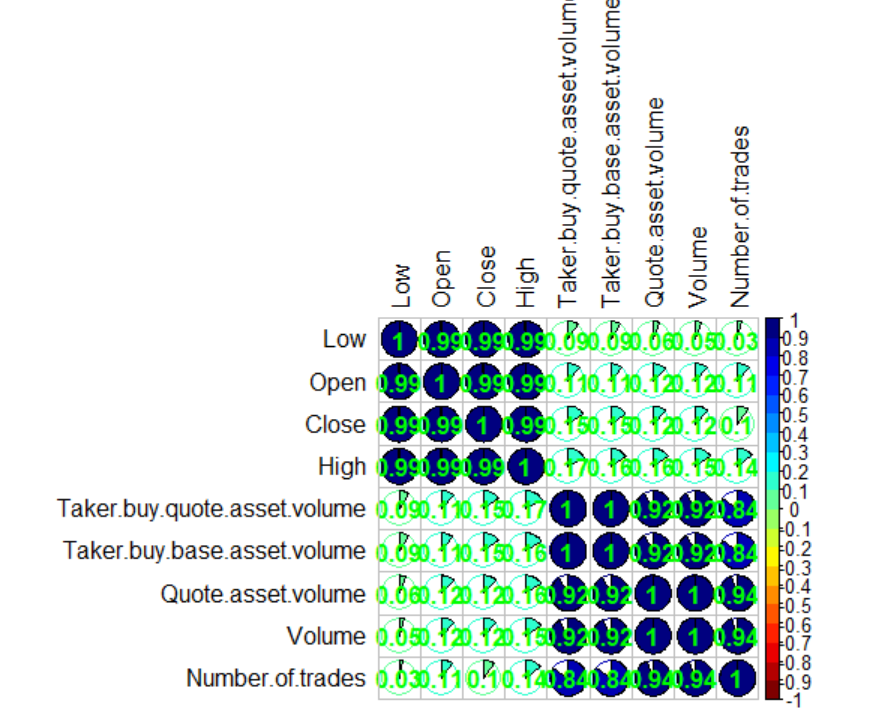
**ДЗ №8.**

**Часть 1. Изучено.**

**Часть 2. Проверка наличия мультиколлинеарности в собственных данных с помощью корреляционной матрицы и фактора инфляции дисперсии.**

*Код:*

Изображение выглядит как текст

Автоматически созданное описание*График:*

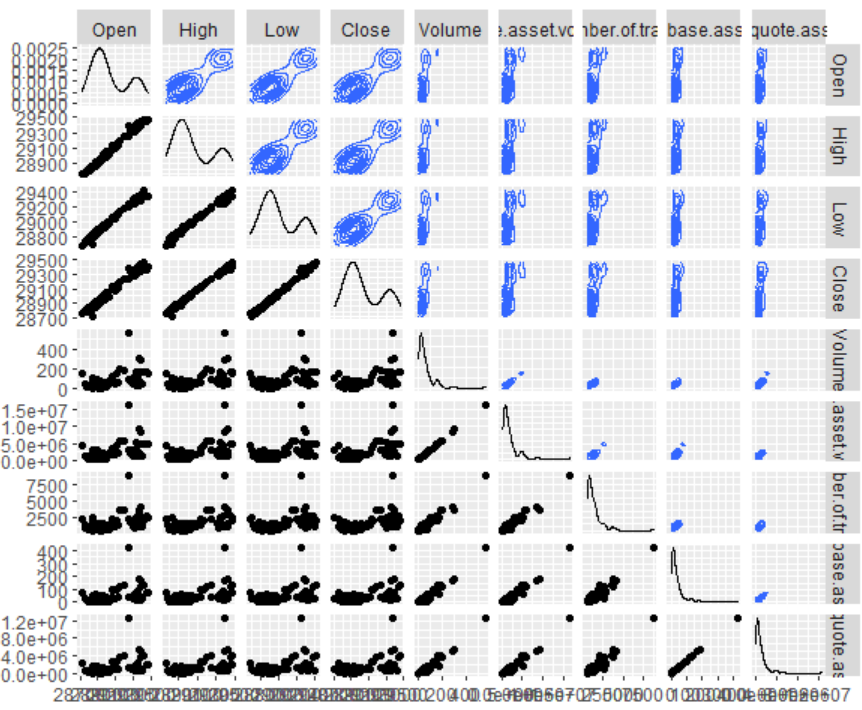
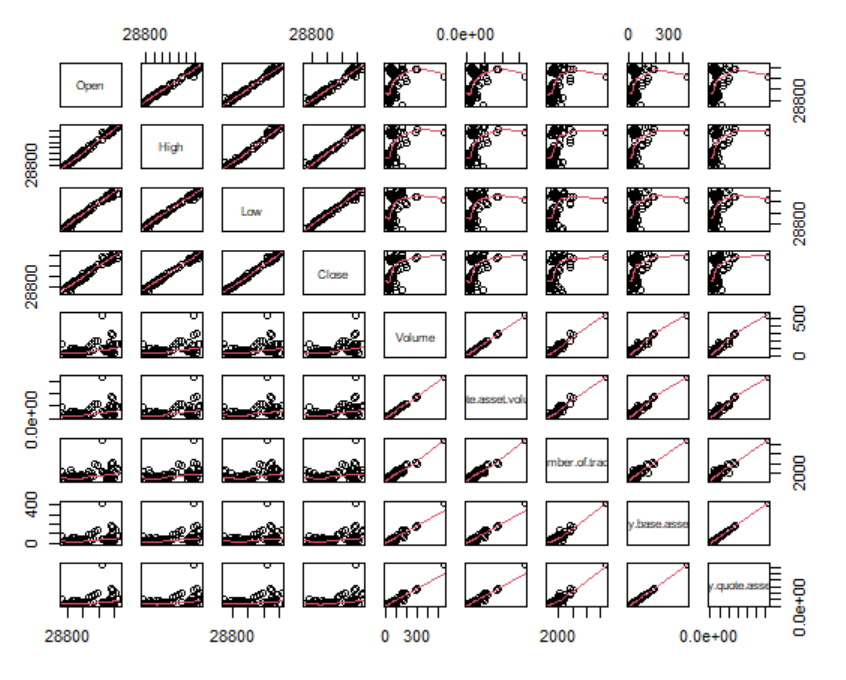
*По графику видно, что столбцы делятся на 2 группы эквивалентной коллинеарности. Можно сделать вывод о том, что в моих данных есть только 2 линейно независимых столбца.*

*Факторы инфляции дисперсии подтверждают мои слова:*Изображение выглядит как текст

Автоматически созданное описание*Обычно предельными значениями берут 5(реже 10). В случае моих данных числа настолько огромные, что наличие линейной связи очевидно.*

**Часть 3. Изучение формы связи между переменными.**

*Код:(Непрерывные данные)*

Изображение выглядит как текст

Автоматически созданное описание*Графики:*

*Характер связи между переменными проще всего выявить, используя соответствующие графические средства.*

*В части 2 мы показали, что столбцы в моих данных поделились на 2 группы, внутри которых они линейно зависимы друг от друга.*

*Этот же эффект мы можем наблюдать на диаграммах рассеяния.*

*Однако, остается вопрос: какая функциональная связь между группами? Рассмотрим зависимость столбцы High от Number.of.trade. Навскидку по графикам(особенно по 2-х мерной ядерной оценке плотностей) можно предположить, что связь имеет константный вид, то есть значение High пробегает от значения 28000 до 30000, а Number.of.trade принимает одни и те же значения.*

*Мои слова подтверждает значение оценки EDF*

Изображение выглядит как текст

Автоматически созданное описание*При значение edf > 1.5 можно делать вывод о нелинейной зависимости данных.*

*Категориальные данные.*

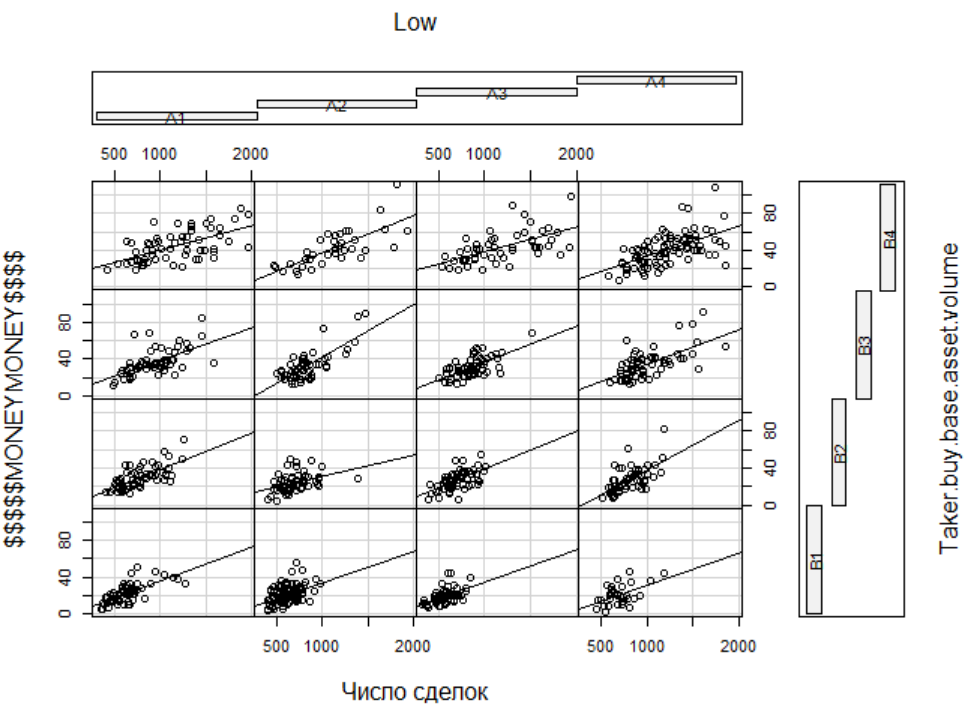
*В моих данных нет категориальных переменных, поэтому я сгенерирую их, разбив 2 непрерывные переменные на квантили.*

*Код:*

Изображение выглядит как текст

Автоматически созданное описание

*График:*

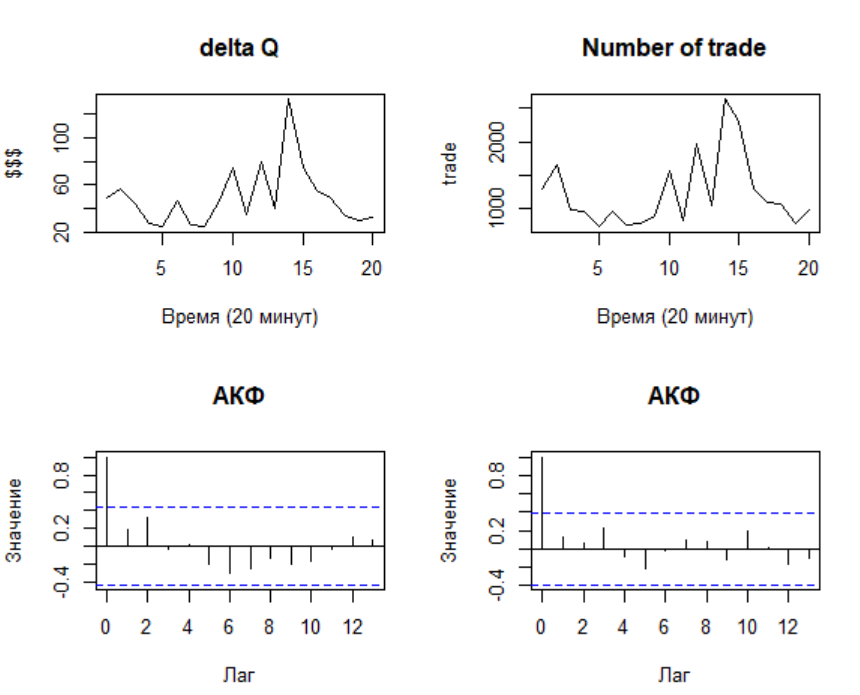
*График получился красивый, но вывод по нему сделать сложно. Можно лишь сказать, что во всех группах имеет место ненулевая корреляция между числом сделок и деньгами.*

**Часть-4. Графики автокорреляционных функций.**

*Код:*

Изображение выглядит как текст

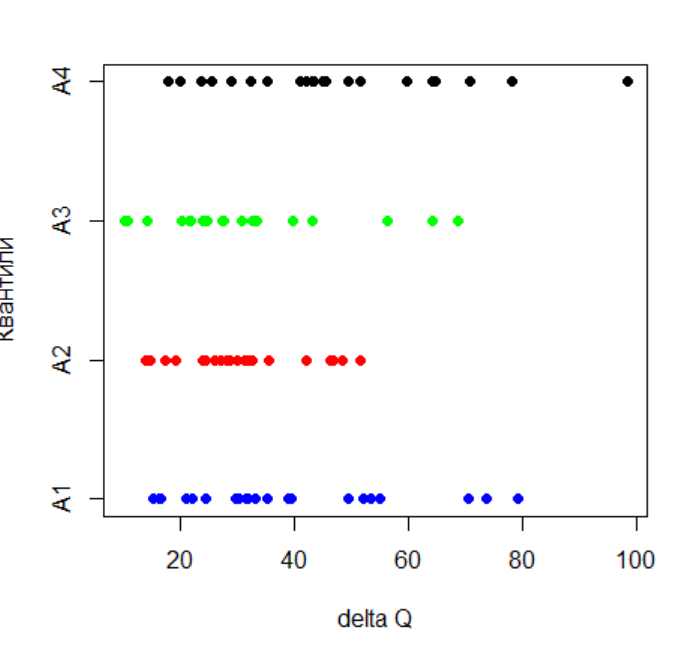
Автоматически созданное описание*Графики:*

*Все данные лежат внутри полосы. Также можно заметить, что на 2 ом графике АКФ имеет синусоидальную форму => исходный график имеет некоторую синусоидальную зависимость.*

***Часть 5. Дисперсионный анализ.***

Изображение выглядит как текст

Автоматически созданное описание*Код:*

*По графику можно сделать вывод о том, что средние значения очень близки друг к другу.*

Изображение выглядит как текст, экран, закрыть

Автоматически созданное описание*Однако p-value очень маленькое, значит нам нужно отвергнуть гипотезу о равенстве средних значений.*

*Для двухфакторного:*

Изображение выглядит как текст

Автоматически созданное описание*Все p-value очень маленькие. Мы отвергаем нулевые гипотезы.*

*Вывод: данные и факторы существенно зависят.*