Министерство образования Республики Беларусь

Учреждение образования

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Инженерно-экономический факультет

Кафедра экономической информатики

**ОТЧЕТ**

по лабораторной работе №7

по предмету «Системы и технологии интеллектуальной обработки данных»

Применение инструментов Data Mining

|  |  |
| --- | --- |
| Выполнил: | Студент группы 772303  Котковец А.С. |
| Проверил:    Минск, 2020 | Ассистент кафедры ЭИ Кунцевич А.А. |

**Задания**

В соответствии с индивидуальным вариантом задания построить  мродель линейной регресиии и выполнить её оценку

**Выполнение**

1.Указыаем исходные данные. Каждая из пяти пар состоит из года и средней процентной ставки:



 2. Взглянем на данные. Сначала построим точечный график, используя наши данные и замечаем, что он выглядит линейным. Чтобы подтвердить наши подозрения, мы найдем корреляцию между годом и средней процентной ставки:

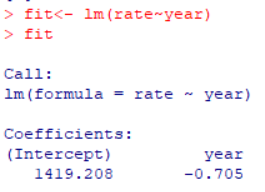
3. Прежде чем найти регрессию наименьших квадратов, надо решить, что будет пояснительной переменной и что будет переменной отклика. В нашем случае правильнее будет выбрать пояснительной переменной год, а переменной отклика процентную ставку. Это было выбрано потому, что процентная ставка может изменяться во времени, а не время изменяется по мере изменения процентных ставок.

Команда для выполнения регрессии наименьших квадратов является команда *lm.*

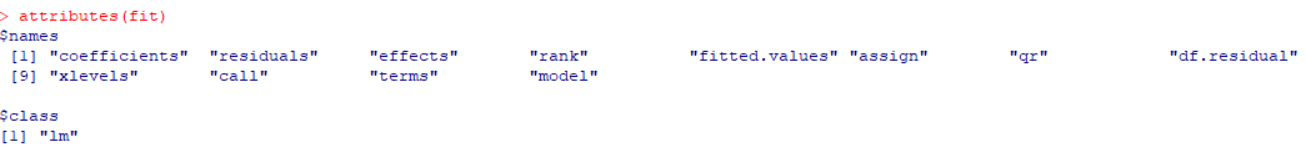
Процентная ставка будет являться переменной отклика, а год поясняющей переменной, то это означает, что линейная регрессии может быть записана в виде  slope-intercept :

rate=(slope)year+(intercept)

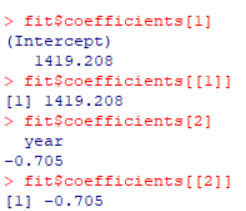
Эта связь определяется в команде *lm,* вы пишете вектор, содержащий переменную отклика, тильда ("~"), и еще одни вектор, содержащий объясняющую переменную:



4. Функция lm возвращает переменную с большим количеством информации. Если узнать, что еще хранится в переменной, можно использовать команду attributes:



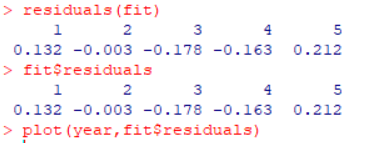
5. Можно увидеть переменные коэффициенты в fit



6. Оценка процентной ставки в 2020 году.



7. Вычисление остатков



8. Тизер для тех видов анализов.

