Федеральное агентство по образованию

Государственное образовательное учреждение

высшего профессионального образования

«Южно-Уральский государственный университет»

Кафедра "Прикладной математики"

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №1

по дисциплине «Эконометрика»

Тема: «Множественная регрессия»

Выполнил:

Проверил: Литвинов А.О.

«\_\_\_» \_\_\_\_\_\_\_\_\_ 2016 г.

Челябинск - 2016

**library("dplyr")**

**library("ggplot2")**

**library("GGally")**

**library("psych")**

*# 1. Постройте матрицу парных коэффициентов корреляции. Установите какие факторы мультиколлинеарны.*

*# 2. Постройте уравнение множественной регрессии в линейной форме с полным набором факторов.*

*# 3. Оцените статистическую значимость уравнения регрессии и его параметров с помощью критериев Фишера и Стьюдента.*

*# Отберите информативные факторы . Постройте уравнение со статистически значимыми факторами*

*# Оцените качество уравнения через среднюю ошибку аппроксимации*

X = **data.frame**(y = c(39, 68.4,34.8,39,54.7,74.7,71.7,74.5,137.7,40,53,86), #Валовой доход за год

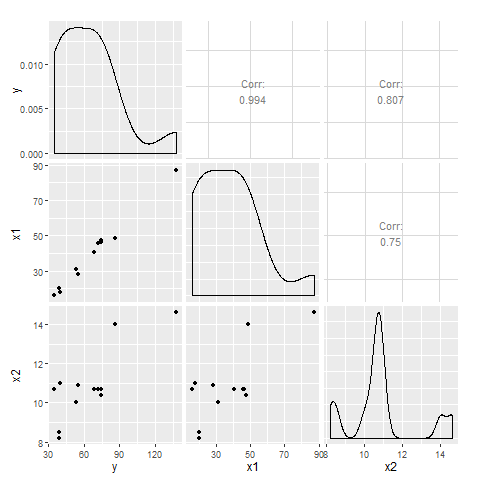
x1 = c(20,40.5,16,20,28,46.3,45.9,47.5,87.2,17.7,31.1,48.7), #Основных фондов

x2 = c(8.2,10.7,10.7,8.5,10.9,10.7,10.7,10.4,14.6,11,10,14)) #Оборотных средств

**png**(filename= "1.png")

**ggpairs**(X)

**dev.off**()



linearModel\_x1\_x2 <- **lm**(y~x1 + x2, data = X)

linearModel\_x1 <- **lm**(y~x1, data = X)

linearModel\_x2 <- **lm**(y~x2, data = X)

result <- **data.frame**(y = X$y,

x1\_x2 = **predict**(linearModel\_x1\_x2),

x1 = **predict**(linearModel\_x1),

x2 = **predict**(linearModel\_x2))

*# предполагается, что выборки должны быть нормальными, чего нет, что с этим делать?*

*###############################проверка равенства средних############################*

resultTtest <- **data.frame**(p\_X1\_X2 = **t.test**(result$y, result$x1\_x2)$**p.value**,

p\_X1 = **t.test**(result$y, result$x1)$**p.value**,

p\_X2 = **t.test**(result$y, result$x2)$**p.value**)

*# p\_X1\_X2 | p\_X1 | p\_X2*

*# 1 | 1 | 1*

*# можно с уверенностью сказать, что среднии равны, что на самом деле так - они все равны 64.45833*

*#####################################################################################*

*###############################проверка равенства дисперсий##########################*

resultFishertest <- **data.frame**(p\_X1\_X2 = **var.test**(result$y, result$x1\_x2)$**p.value**,

p\_X1 = **var.test**(result$y, result$x1)$**p.value**,

p\_X2 = **var.test**(result$y, result$x2)$**p.value**)

*# p\_X1\_X2 | p\_X1 | p\_X2*

*# 0.994 | 0.983 | 0.489*

*# можно сделать вывод, что дисперсии достаточно схожие с исходными, кроме модели, построенной на основании второй переменной - x2, причину этого можно увидеть, если посмотреть на гистограмму распределения признака - она координально отличается от y и x1*

*# 95 confidence interval отношений дисперсий*

*# y~x1 0.2916438 3.5191452*

*# y~x2 0.4416783 5.3295486*

*# y~x1\_x2 0.2890907 3.4883381*

#####################################################################################

sd\_x1\_x2 <- **sd**(result$y - result$x1\_x2)

sd\_x1 <- **sd**(result$y - result$x1)

sd\_x2 <- **sd**(result$y - result$x2)

*# model\_x1\_x2 - 1.863*

*# model\_x1 - 3.269*

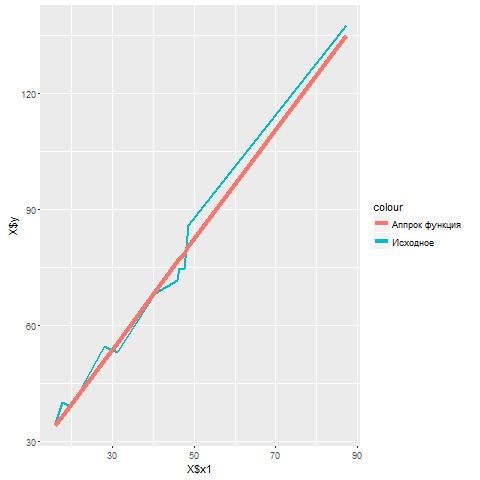
*# model\_x2 - 16.975*

*#****summary****(linearModel\_x1\_x2)*

*#****summary****(linearModel\_x1)*

*#****summary****(linearModel\_x2)*

**Выводы:** на основании парной коррелограмы, мы можем сделать вывод о том, что признак x1 и х2 сильно скоррелированы, следовательно один из них мы отбросим. Также мы видим, что корреляция с целевой переменной у х1 больше, чем у х2, следовательно выбросим х2. Итого наша итоговая модель приобретет вид: **y = 11.30509 + 1.42089 \* x1.** Значение среднеквадратичного отклонения от предсказываемого равно **3.26903**, что вполне приемлемо, для предсказываемого значения с матожиданием **64.45833** и среднеквадратичного отклонения равного **28.7666**

****