Инструкция по выполнению 2 этапа проекта.

После того, как данные собраны, необходимо их проанализировать.

Факторы, которые мы исследуем, имеют разную природу и обладают свойствами, которые нам могут помешать. Так, например, **величина заработной платы** обладает свойством сезонности, а значит, **необходимо** провести корректировку этого ряда (о том, как это сделать написано ниже).

Начать наш анализ можно с построения графиков. Сперва построим графики для каждой переменной отдельно. Затем построим диаграммы рассеяния (scatter plot) для зависимой и независимых переменных.

Визуальный анализ может указать нам на некоторые свойства временного ряда. Например, можно увидеть тренд или сезонность. График типа scatter plot может дать информацию о форме зависимости переменных, что необходимо будет учесть при построении регрессионной модели.

# График

Чтобы построить график нужно нажать на меню Quick и выбрать Graph:

Затем выбрать необходимые переменные, для которых хотите построить график:



Далее откроется меню, где можно выбрать тип графика и настроить его вид:



Получаем график:



Каждый график обязательно нужно подписать и дать его описание – что на нём видите, какие выводы можно сделать, предложения по дальнейшему исследованию и т.п.

# Сезонность

Чтобы провести корректировку на сезонность нужно в окне соответствующей переменной выбрать в меню пункт PROC и перейти к опции Seasonal Adjustment, после чего выпадет дополнительное меню с разными методами корректировки (Census X13, Census X12, X11

(Historical), Tramo/Seats, Moving Average Methods).):



# Стационарность.

Важным свойством временных рядов является стационарность, благодаря которой ряды не меняют своих вероятностных свойств с течением времени. Для проверки на стационарность можно воспользоваться тестом Дики — Фуллера. Если ряд не стационарен, то необходимо привести его к стационарному виду путём преобразований, например, взяв первые разности. Далее следует повторить проверку на стационарность.

# Тест Дики — Фуллера

Для проведения теста Дики-Фуллера (или другого на единичные корни) необходимо в окне временного ряда (переменной) выбрать в меню пункт View и в открывшемся меню выбрать Unit Root Test:



У вас откроется меню, где вы можете выбрать необходимые параметры:



Окно с результатами теста:



В работе обязательно опишите результаты теста словами и ваше предложения по дальнейшей работе.

# Корреляции

В рамках второго этапа вам также предлагается построить корреляционные матрицы.

Чтобы построить матрицу корреляций выберите перейдите в меню Quick подменю group statistics и выберете команду correlations.



В появившемся окне выберете необходимые переменные.



Выделите значения, которые посчитаете значимыми, сделайте выводы.

*Результат проведённого анализа необходимо представить в текстовом документе формата word. Названия файла должно быть таким: «ххх.анализ данных», где xxx – название региона. Например, «Красноярский край. Анализ данных.docx». Высылать на почту* *ptukpetov@hse.ru* *с темой письма «xxx.Этап 2», где xxx – название региона. Например, «Республика Тыва. Этап 2».*