# Введение

Проблемы связанные с рождаемостью, а соответственно и с коэффициентом фертильности выражающим ее, актуальны сейчас во всем мире, как никогда. В ряде стран — это проблема старения населения и как следствие увеличивающиеся нагрузка на экономически активное население, пенсионную систему, государственный бюджет в целом, в ряде других стран — это нехватка ресурсов — голод, бедность, вызванная перенаселением, сохранением высоких уровней рождаемости.

Для того, чтобы понять можем ли мы решить эти проблемы и если да, то как нам прежде всего нужно проанализировать от чего зависит уровень рождаемости. Именно это является целью моей работы.

1. **СБОР ДАННЫХ И ОТБОР ФАКТОРОВ**

Для проведения исследования были отобраны следующие факторы, влияющие на коэффициент фертильности.

К факторам относятся:

*X1* – ВВП на душу населения ( с поправкой на паритет покупательской способности). CIA Factbook 2008

*X2* – выбросы CO2 на душу населения. World Resource Institute 2003

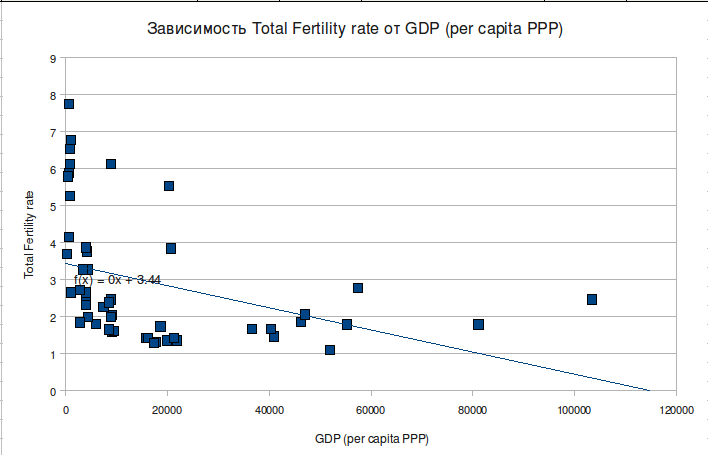
*X3* – индекс демократии страны. The Economist 2008

*X4* – потребление алкоголя на душу населения. OECD Health Data 2005

В качестве результирующего признака (коэффициент фертильности) были использованы данные из CIA Factbook 2008.

Для обеспечения репрезентативности выборки, отбор стран проводился путем деления всей совокупности (рейтинг стран по ВВП на душу населения) на 5 равных частей и отбора из них 10 стран наиболее близких к внутригрупповой арифметической средней.

1. **ИССЛЕДОВАНИЕ ВЛИЯНИЯ ОТДЕЛЬНЫХ ФАКТОРОВ** 
   1. **Исследование влияния ВВП на душу населения (с поправкой на паритет покупательской способности) на коэффициент фертильности.**





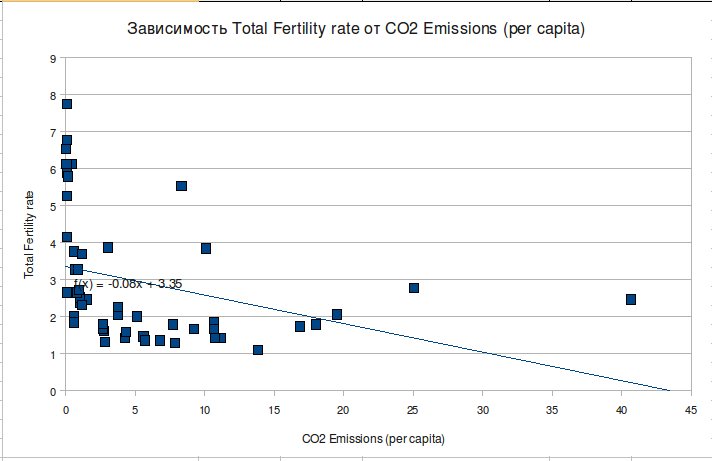
При рассмотрении поля корреляции можно сделать следующие выводы:

* зависимость проявляется лишь со значительным увеличением ВВП примерно с 10000$, а при сравнительно низких его значениях представляет из себя «белый шум» т.е. отсутствие зависимости.
* линейная модель является наиболее подходящей, т.к. она проста и более легка в интерпретации.

По коэффициенту корреляции связь умеренная. Вариация коэффициента фертильности лишь на 15% объясняется вариацией ВВП, а остальные 85% другими неучтенными в этой модели факторами. Средняя ошибка аппроксимации составляет 52%, что намного превышает рекомендуемый 10% уровень. При этом и коэффициенты и в целом построенное уравнение регрессии статистически значимо.

Несмотря на неоднозначность проведенного исследования X1, как один из факторов влияющих на фертильность, заслуживает включения во второй этап - исследование совокупности факторов.

* 1. **Исследование влияния выбросов CO2 (на душу населения) на коэффициент фертильности.**





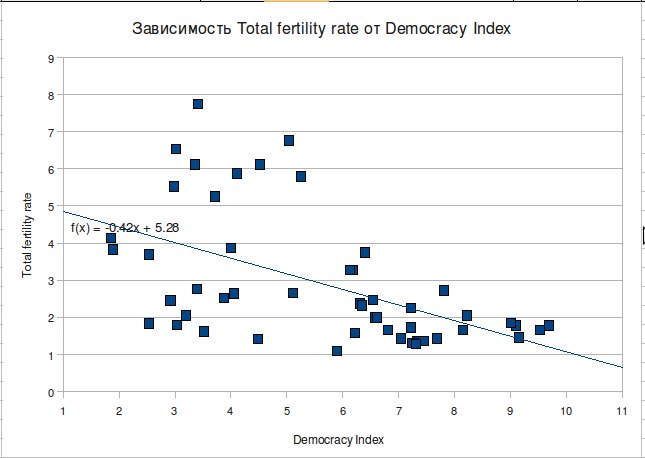
При рассмотрении поля корреляции можно сделать следующие выводы:

* зависимость проявляется лишь со значительным увеличением выбросов CO2 примерно с 3 тысяч тонн CO2 на 1000 человек населения, а при сравнительно низких значениях выбросов представляет из себя «белый шум» т.е. отсутствие зависимости.
* линейная модель является наиболее подходящей, т.к. она проста и более легка в интерпретации.

По коэффициенту корреляции связь умеренная. Вариация коэффициента фертильности лишь на 12% объясняется вариацией выбросов СО2, а остальные 88% другими неучтенными в этой модели факторами. Средняя ошибка аппроксимации составляет 54%, что намного превышает рекомендуемый 10% уровень. При этом и коэффициенты и в целом построенное уравнение регрессии статистически значимо.

Несмотря на неоднозначность проведенного исследования X2, как один из факторов влияющих на фертильность, заслуживает включения во второй этап - исследование совокупности факторов.

* 1. **Исследование индекса развития демократии на коэффициент фертильности.**



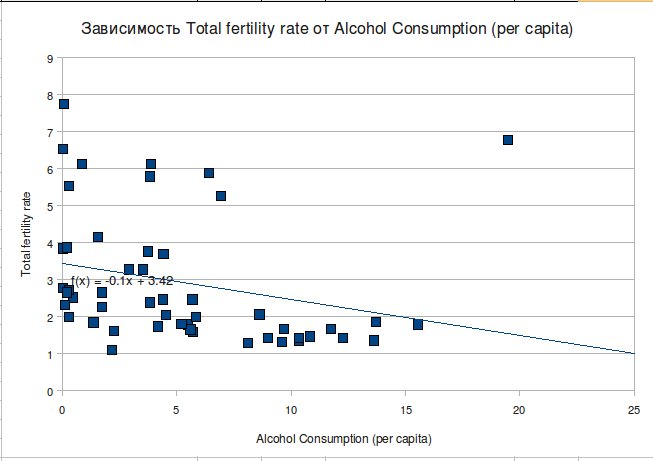


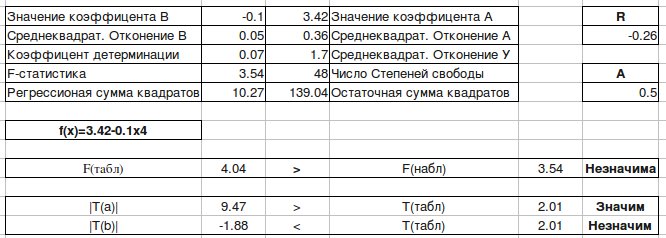
При рассмотрении поля корреляции можно сделать следующий вывод, что линейная модель является наиболее подходящей, т.к. она наиболее проста и визуально достаточно точно отражает зависимость Y от фактора X3.

По коэффициенту корреляции связь заметная. Вариация коэффициента фертильности на 28% объясняется вариацией уровня демократии, а остальные 72% другими неучтенными в этой модели факторами. Средняя ошибка аппроксимации составляет 44%, что намного превышает рекомендуемый 10% уровень. При этом и коэффициенты и в целом построенное уравнение регрессии статистически значимо.

Проведенный анализ показал, что фактор X3, как один из факторов влияющих на фертильность, заслуживает включения во второй этап - исследование совокупности факторов.

* 1. **Исследование влияния потребление алкоголя (на душу населения) на коэффициент фертильности.**





При рассмотрении поля корреляции можно сделать следующие выводы:

* зависимость проявляется лишь со значительным увеличением потребления алкоголя примерно с 3 литров на душу населения, а при сравнительно низких его значениях представляет из себя «белый шум» т.е. отсутствие зависимости.
* линейная модель является наиболее подходящей, т.к. она проста и более легка в интерпретации.

По коэффициенту корреляции связь слабая. Вариация коэффициента фертильности лишь на 7% объясняется вариацией потребления алкоголя, а остальные 93% другими неучтенными в этой модели факторами. Средняя ошибка аппроксимации составляет 50%, что намного превышает рекомендуемый 10% уровень. При этом и коэффициент b и в целом построенное уравнение регрессии статистически незначимо.

Проведенный анализ показал, что включение фактора X4 во второй этап исследование совокупности факторов неоправданно.

**3. ИССЛЕДОВАНИЕ ВЛИЯНИЯ СОВОКУПНОСТИ ФАКТОРОВ**

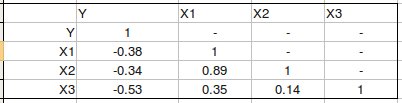
В качестве факторов, оказывающих влияние на коэффициент фертильности, после предварительного исследования были отобраны только три:

*X1* – ВВП на душу населения ( с поправкой на паритет покупательской способности). CIA Factbook 2008

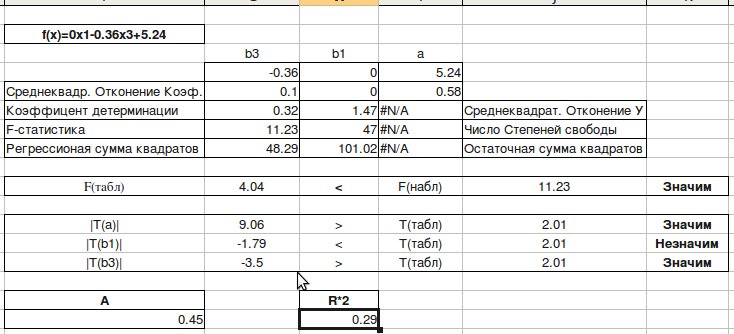
*X2* – выбросы CO2 на душу населения. World Resource Institute 2003

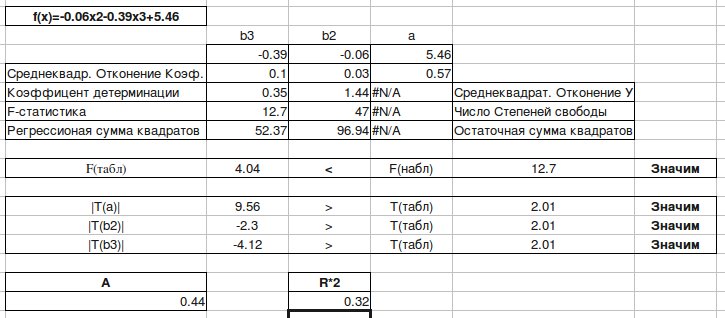
*X3* – индекс демократии страны. The Economist 2008

Рассчитанные парные коэффициенты корреляции представим в виде корреляционной матрицы.

Из анализируемых данных видно, что наблюдается высокая связь между факторами Х1 и Х2, означающая нецелесообразность построения на их основе модели множественной регрессии.

Из анализируемых факторов наибольшее влияние на фертильность оказывает третий фактор – индекс развития демократии. Так как он наилучшим образом коррелирует с результативным признаком. Поэтому возьмём его в качестве базового для построения 2 моделей множественной линейной регрессии.

* У от Х1 и Х3
* У от Х2 и Х3



Из приведенного выше анализа видно, что множественная линейная регрессия У от Х2 и Х3 более отвечает задачам оценки коэффициента фертильности по сравнению с множественной линейной регрессией У от X1 и Х3.

Теперь произведем сравнение выбранной модели множественной регрессии с моделью парной регрессии зависимости коэффициента фертильности от индекса демократии. Сравнив показатели обоих, можно сделать вывод, что множественная регрессия превосходит парную по скорректированному и простому коэффициенту детерминации на 6% и 7% соответственно. При этом cредняя ошибка аппроксимации у обоих моделей примерно равны. Несмотря на рост скорректированного показателя детерминации, чтобы избежать ошибок при принятии управленческих решений, связанных с трудностями интерпретации множественной модели, следует принять модель зависимости коэффициента фертильности от уровня развития демократии — как рекомендуемую к дальнейшему применению.

# Заключение

В ходе проведённого исследования была изучена зависимость между результирующим признаком и четырьмя факторами как в отдельности, так и в совокупности. Для этого были построено 4 парных моделей линейной регрессии и 2 модели множественной регрессии.

Проведённое исследование показало, что только 1 из 6 моделей пригодна для практического применения и интерпретации. Зависимость коэффициента фертильности от индекса развития демократии.

Y=**5.28-0.42x3**

Изменение в уровне демократии на 28% объясняет вариацию в коэффициенте фертильности.

Несмотря даже на высокую субъективность составления индекса демократии журналом The Economist, в целом зависимость можно считать значительной.

Складывается впечатление, что социальная система устроенна удивительно разумно, с развитием общества ресурсы становятся все более дефицитными. Показательны вычисления ученых подсчитавших, что для поддержания уровня потребления современной Калифорнии для всего человечества примерно 6 млрд. потребуется ресурсы шести планет Земля. Таким образом с развитием общества, и соответственно демократии, как наиболее совершенной из известных форм социальной организации падает и рождаемость с целью сохранения ресурсов. Ведь большее количество детей в семье означает при прочих равных меньший уровень потребления для взрослых, и зачастую для будущих родителей на первом месте стоит именно второе, таким образом осуществляется саморегуляция социальной системы в плане рождаемости.

*Приложение А*

**Данные**

