***Анализ данных по суицидам за период с 1985 по 2016 годы.***

Исследование на основе датасета "Suicide\_rates", включающее в себя различные методы анализа данных, в том числе: построение корреляционной матрицы, гистограмм, парных графиков, а также анализ временных рядов (различных промежутков) и демографических характеристик (+социальных). В этом документе представлены основные результаты исследования, в которых мною были выделены факторы, влияющие на статистику по самоубийствам в различных ключах.

Какие аналитические работы были проведены и с какой целью (в конце предоставлен короткий ИТОГ, содержащий гипотезу, которая была выдвинута в ходе работы и ее подтверждение; все таблицы можно найти в Jupiter notebook):

**Корреляционный анализ:**

В первую очередь, была построена корреляционная матрица, которая показала взаимосвязи между различными факторами. Коэффициенты корреляции, близкие к 1, указывают на сильную зависимость между факторами. В данном анализе были выявлены следующие сильные зависимости:

* HDI for year и gdp\_per\_capita ($) с коэффициентом корреляции 0,77.
* Суициды и численность населения с коэффициентом 0,62.
* suicides/100k pop и suicides\_population\_ratio с коэффициентом 1 я не стала рассматривать, так как это одно и то же.

**Гистограммы:**

Для лучшего понимания распределения данных были построены гистограммы по ключевым переменным. При этом было опробовано несколько значений размера "корзины" (30, 40, 100, 200), и оптимальным на мой взгляд оказалось значение 80. Это значение позволило получить достаточную детализацию без потери важных особенностей данных.

**Парные графики:**

Далее были рассмотрены парные графики, которые позволили визуализировать взаимосвязи между парами переменных. Анализ показал минимальное количество выбросов, что свидетельствует о надежности данных. Однако для более детальных выводов потребовался дальнейший анализ взаимосвязей параметров.

**Топ-10 стран по количеству суицидов:**

Был проведен количественный анализ, который выявил топ-10 стран (топ -в контексте набольших показателей) по абсолютному количеству суицидов. Этот анализ не учитывал процентное соотношение к населению, поэтому нельзя провести достоверный сравнительный анализ. В дальнейшем для исправления такой погрешности были рассмотрены данные в процентном соотношении, что изменило картину, но РФ осталась в первой тройке. Обе гистограммы преимущественно указывают, что большая часть суицидов приходится на развитые или европейские страны.

**Распределение суицидов по временным периодам (по годам):**

Анализ распределения суицидов по годам показал, что наибольшее число случаев наблюдалось в период с 1995 по 2005 годы. Это может быть связано с различными социально-экономическими факторами, характерными для данного периода. Но этот вопрос требует детального изучения с исторической, политической, экономической и социальных точек зрения.

**Возрастные и половые группы:**

Исследование распределения суицидов среди возрастных и половых групп показало, что наибольшее число самоубийств совершают люди в возрасте 35-54 лет, преимущественно мужчины. Это поколение, известное как "boomers", было наиболее подвержено суициду в период с 1985 по 2016 годы.

**Половые различия по странам:**

Для более детального анализа были изучены данные по странам, чтобы определить, в каких из них суициды чаще совершают мужчины или женщины. Оказалось, что мужчины чаще совершают суицид в крупных развитых странах, что может быть связано с большей экономической и административной ответственностью и психологическим давлением. Анализ женщин-суицидниц показал, что причины могут быть связаны с тяжелыми социальными ролями и необходимостью совмещения семейных и профессиональных обязанностей, а также их исторической социальной роли в обществе, но этот вопрос также требует более детального рассмотрения.

**Процентное соотношение мужчин и женщин по странам:**

Далее был проведен анализ процентного соотношения мужчин и женщин, совершающих суициды в каждой стране. Рассмотрены данные о том, какое количество суицидов приходится на мужчин и женщин от общего числа самоубийств. Это позволило выделить 10 стран с наибольшими показателями и предположить, что высокий процент суицидов среди женщин связан с культурно-психологическим давлением.

**Динамика суицидов по годам и полу:**

Анализ динамики суицидов по годам и полу показал, что количество мужских самоубийств имеет тенденцию к резкому изменению, в то время как показатели среди женщин остаются примерно стабильными.

**Влияние экономических факторов:**

Был проведен анализ влияния экономических факторов на количество суицидов. Выяснилось, что наибольшее количество суицидов наблюдалось в переходные периоды между веками, что может быть связано с изменениями в экономическом развитии. Зависимость количества суицидов от ВВП на душу населения подтверждает связь с экономическими трудностями. Но опять-таки для более точных умозаключений потребуется дополнительный анализ конкретно по этому фактору.

**Влияние уровня интеллектуального развития:**

Наконец, был проанализирован уровень интеллектуального развития (HDI) и его влияние на склонность к суициду. Результаты показали, что чем выше индекс развития у человека, тем больше вероятность суицида. Мужчины с индексом развития равного среднему значению и выше среднего совершают суициды чаще, аналогичная тенденция наблюдается среди женщин, только мы наблюдаем меньшее количество женщин, совершивших суицид, со средним показателем

**ИТОГ:**

* *Гипотеза 1*

Высокий уровень HDI и ВВП на душу населения коррелирует с увеличением количества суицидов.

Подтверждение: Корреляция между HDI и ВВП (0,77).

* *Гипотеза 2*

Развитые страны имеют высокий уровень суицидов.

Подтверждение: Гистограммы и процентные соотношения показывают, что развитые страны часто находятся в топе по количеству суицидов.

* *Гипотеза 3*

Мужчины совершают суициды чаще, чем женщины, особенно в экономически развитых странах.

Подтверждение*:* Анализ возрастных и половых групп.

* *Гипотеза 4*

Переходные периоды между веками сопровождаются ростом числа суицидов.

Подтверждение: Анализ суицидов по годам и ВВП на душу населения.

* *Гипотеза 5*

Высокий индекс развития (HDI) у человека связан с большей склонностью к суициду.

Подтверждение: Анализ зависимости суицидов от индекса развития.

* *Гипотеза 6*

Социальное и экономическое давление в развитых странах является основным фактором суицидов среди женщин.

Подтверждение: Процентное соотношение и культурно-психологические факторы.