# Титульный лист

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ**

**ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ**

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет Международный бакалавриат по бизнесу и экономике

Каратуева Анна БМБ2208

Самохвалова Вера БМБ2208

Скублова Яна БМБ2208

Никифорова Мария БМБ2208

**“Анализ привлекательности стран для**

**стартапов”**

# 

Санкт-Петербург 2025

# **Введение**

Стартапы сегодня играют важную роль в экономике, потому что они создают новые рабочие места и создают новые, инновационные решения. Для предпринимателей стоит один из самых важных вопросов: в какой стране лучше открыть свой бизнес?

Цель нашего проекта состоит в том, чтобы выяснить, какие факторы делают страну более привлекательной для стартапов. Доступ к венчурному финансированию, уровень доходов, инфляция, корпоративные налоги и экономическая стабильность — это основные показатели, которые мы выделили на основе изучения различных источников.

Мы создали индекс бизнес-привлекательности, собрав и проанализировав данные по более чем 60 странам. Затем мы проверили, какие факторы действительно влияют на количество стартапов с помощью алгоритма машинного обучения Random Forest.

# 

# **Исследовательская задача**

В нашем исследовании мы хотим выяснить, какие экономические переменные оказывают решающее влияние на привлекательность страны для стартапов. Мы хотим определить, какие страны предлагают лучшие условия для бизнеса, анализируя венчурные инвестиции, ВВП на душу населения, инфляцию, корпоративные налоги, уровень безработицы, среднюю заработную плату.

Мы разработали следующие гипотезы для выполнения задачи:

Гипотеза 1**.** Страны с более высоким ВВП на душу населения более привлекательные для стартапов.

Гипотеза 2**.** Объем венчурных инвестиций является значимым фактором, влияющим на успешность стартапов.

Мы использовали регрессионный анализ и алгоритм машинного обучения Random Forest для проверки этих гипотез и выявления наиболее важных показателей.

# **Обзор литературы**

1. ПЕ, Р., & ЕИ, Г. (2022). Стартап как форма ведения бизнеса. КиберЛенинка.

В этой статье рассматривается стартап как форма современного инновационного бизнеса, анализируются понятия «бизнес» и «предпринимательство», а также особенности стартапов в различных странах.

Эта статья стала основой для выбора экономических показателей (ставки корпоративных налогов, инфляция, ВВП на душу населения и т.д.), которые мы использовали в проекте.

2. Едигаров, Д. (2024). Где лучше всего развивать свой стартап? Сравниваем Европу и США. Хабр.

В статье проводится сравнительный анализ стартап-экосистем Европы и США, рассматриваются преимущества и недостатки каждой из них.

С помощью данной статьи, мы выделили важность наличия развитой предпринимательской экосистемы и венчурного капитала. В нашем исследовании именно страны с развитым венчурным рынком продемонстрировали наиболее высокие показатели привлекательности для стартапов.

3. Алексеевич, С. П. (2022). Рекомендации по развитию стартапа с учетом значимых факторов. КиберЛенинка.

Авторы анализируют ключевые факторы, влияющие на успех стартапов, и предлагают рекомендации по их развитию, основываясь на материалах из научных журналов, индексируемых в Scopus и WoS.

Эта статья помогла нам выбрать показатели в нашей модели и методологии. После прочтения мы поняли, почему умеренные налоги и доступ к венчурному капиталу сильно влияют на успешность стартапов. Также, дополнительно акцентировали внимание на таких странах как Швейцария.

4. Яндекс. (2023). *Какие факторы влияют на успешность небольших стартапов?* Яндекс.Бизнес.

В этой статье рассматриваются факторы успешности стартапов. С помощью этой работы, мы подтвердили важность венчурного финансирования, дополнительно акцентируя внимание на роли экономической стабильности и доходов населения. Это соответствует нашим выводам о важности высокого ВВП как фактора для успеха стартапа.

5. Факторы, влияющие на успех стартапа. (2022). *Startup Jedi.* Retrieved March 24, 2025

​В статье анализируются ключевые элементы, которые определяют успешность стартапов.

Данная статья дополнительно подтвердила выводы нашего регрессионного анализа, в котором количество венчурных инвестиций оказалось главным значимым фактором, который положительно влияет на количество стартапов.

# 

# **Выбор данных**

На начальном этапе нашего проекта мы собираем данные из различных источников, чтобы создать единый датафрейм, который в дальнейшем будет использоваться для анализа и построения модели. Модель должна оценивать, подходит ли та или иная страна для создания стартапа, основываясь на таких показателях, как:

- ВВП на душу населения

- уровень безработицы

- уровень инфляции

- средняя заработная плата

- корпоративные налоги

- объём венчурных инвестиций

- число зарегистрированных стартапов

Библиотеки, которые мы использовали для сбора и начальной обработки данных:

- pandas для работы с данными

- requests для запросов к API

- numpy для обработки числовых данных

1. Данные из API Всемирного банка (ВВП на душу населения)

* Запрашиваем данные через API Всемирного банка по показателю `NY.GDP.PCAP.CD` (ВВП на душу населения в текущих долларах США) за 2023 год
* Фильтруем данные, оставляя только страны (исключая регионы, у которых нет двухбуквенного кода)

2. Данные об уровне безработицы

* Используем локальный файл API\_SL.UEM.TOTL.NE.ZS\_DS2\_en\_csv\_v2\_76310.xls, предоставляемый Всемирным банком
* Оставляем только столбцы с названием страны и уровнем безработицы за 2023 год
* Переименовываем страны для единообразия (например, "Korea, Rep." → "South Korea")
* Вручную добавляем данные для некоторых стран (например, Китая)

3. Данные об инфляции

* Используем локальный файл global\_inflation\_data.xls
* Оставляем столбцы с названием страны и уровнем инфляции за 2023 год
* Вручную добавляем данные для Китая, Южной Кореи и Словакии

4. Данные о средней заработной плате

* Используем файл Country\_Region\_Income.xlsx
* Вручную добавляем данные для ОАЭ, Венесуэлы, Бразилии, Индии и Индонезии

5. Индекса сложности ведения бизнеса

* Используем файл Global\_Business\_Complexity\_Index\_2023.xlsx
* Переименовываем страны для единообразия (например, "Russia" → "Russian Federation")
* Заменяем отсутствующие значения пустыми строками

6. Данных о корпоративных налогах

* Используем файл 1980\_2023\_Corporate\_Tax\_Rates\_Around\_the\_World\_Tax\_Foundation\_1.xlsx
* Форматируем налоговые ставки в процентный формат.
* Переименовываем страны (например, "United States of America" → "United States")
* Вручную добавляем данные для Ирана

7. Данные о венчурных инвестициях

* Используем файл statistic\_id1480496\_venture\_capital\_funding\_rounds\_worldwide\_2023.xlsx
* Переименовываем столбцы и добавляем данные для России
* Заменяем отсутствующие значения пустыми строками

8. Данные о количестве зарегистрированных стартапов

* Используем файл `Total Startup Output.xlsx`.
* Переименовываем страны (например, "Russia" → "Russian Federation").
* Преобразуем значения в числовой формат.

**Проблемы, возникшие при сборе данных:**

1. Несоответствие названий стран

В разных источниках страны могут называться по-разному (например, "Russia" и "Russian Federation"). Для решения этой проблемы мы вручную переименовываем страны, чтобы названия были единообразными.

2. Отсутствующие данные

Некоторые показатели отсутствуют для определённых стран. Мы либо оставляем их как `NaN`, чтобы в дальнейшем провести необходимые замены уже для итогового датафрейма, либо вручную добавляем недостающие значения, если они доступны из других источников.

3. Разные форматы данных

Например, корпоративные налоги представлены в виде строк с символом %. Мы преобразуем их в единый формат для удобства анализа.

# 

## **Объединение данных в единый датафрейм**

Объединяем все датафреймы (df1 – df8) по столбцу Country с использованием внешнего соединения (outer), чтобы сохранить все страны из всех источников. Сортируем данные по названию страны и сбрасываем индексы.

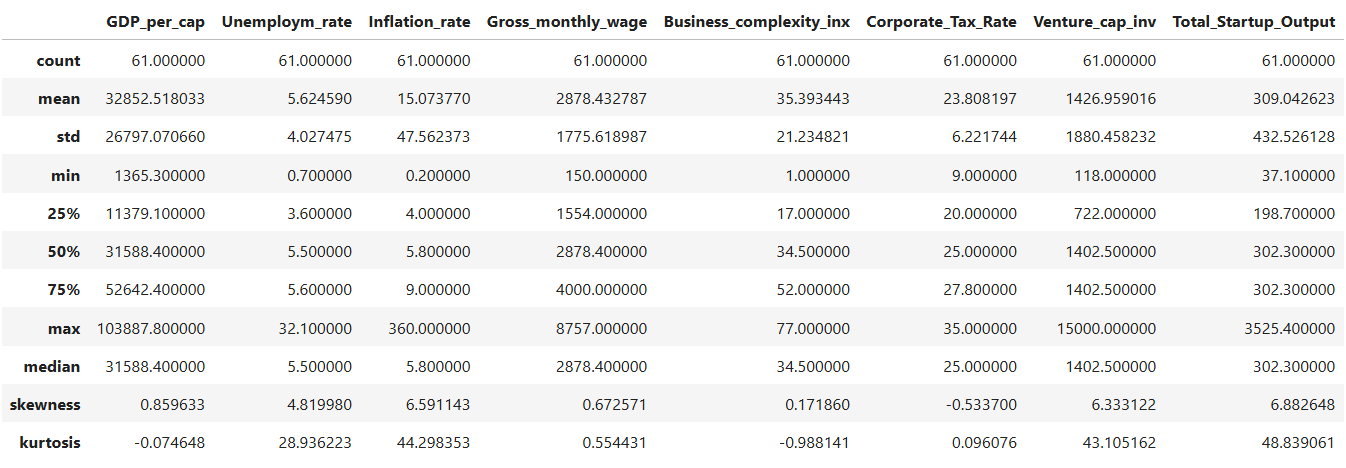
**Датафрейм со всеми странами (df)**: Содержит все страны из исходных данных, но с пропусками во многих показателях.

**Датафрейм с 20 странами (df\_20)**: Содержит только страны, для которых есть данные по всем показателям.

**Датафрейм с 60 странами (df\_final)**: Расширенный набор стран, где пропущенные значения заменены на средние значения по соответствующим столбцам. Используется для построения модели.

## **Разведывательный анализ данных**

1. **Расширенный описательный анализ данных**

****

В первую очередь мы провели расширенный описательный анализ данных, добавляя к стандартной статистике (среднее, стандартное отклонение, минимум, максимум и т.д.) три важные метрики: медиану, асимметрию (skewness) и эксцесс (kurtosis).

Асимметрия измеряет степень отклонения распределения данных от симметричного (нормального) распределения. Эксцесс помогает понять, насколько данные склонны к выбросам. Если эксцесс высокий, это может указывать на наличие выбросов, которые могут повлиять на результаты анализа или моделирования.

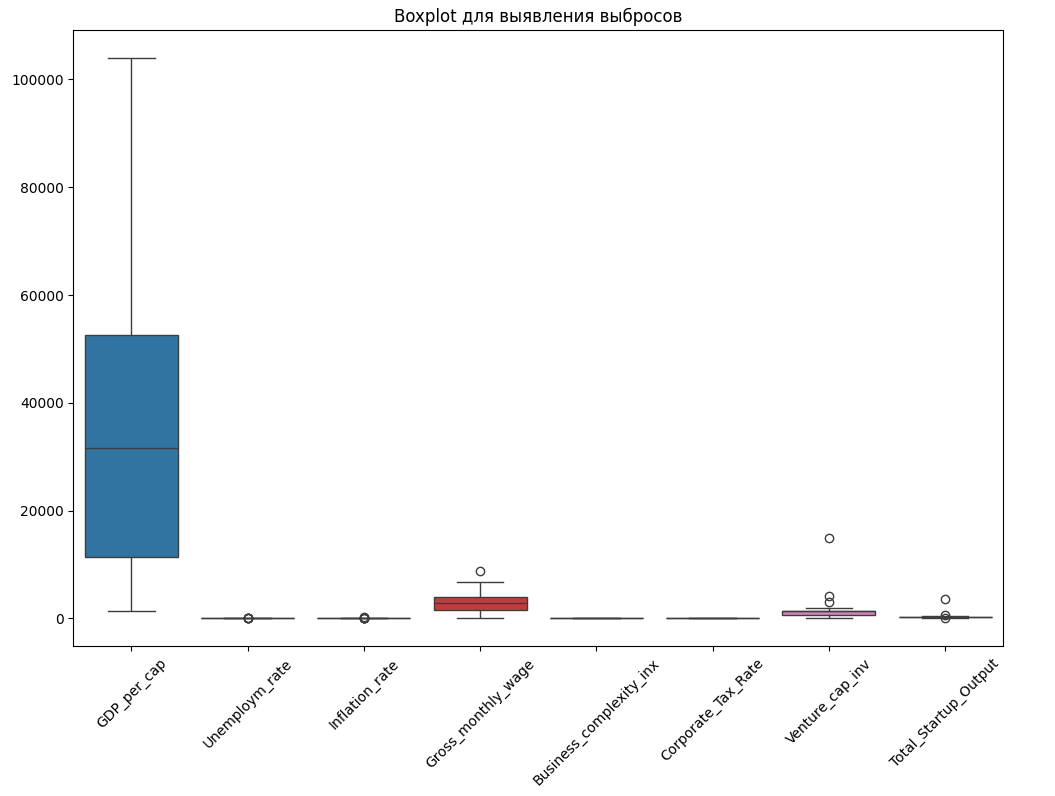
## 

## **Интерпретация результатов описательной статистики**

Распределение ВВП на душу населения имеет небольшую положительную асимметрию (0.86), что указывает на наличие стран с высоким ВВП, при этом эксцесс близок к нулю (-0.07), что говорит о близости к нормальному распределению с небольшим уплощением. Уровень безработицы и инфляция демонстрируют высокую положительную асимметрию (4.82 и 6.59 соответственно) и эксцесс (28.94 и 44.30), что указывает на наличие выбросов — стран с экстремально высокими значениями этих показателей (например, Венесуэла). Средняя зарплата имеет небольшую положительную асимметрию (0.67) и эксцесс близкий к нулю (0.55), что свидетельствует о нормальном распределении. Индекс сложности ведения бизнеса показывает равномерное распределение (асимметрия 0.17, эксцесс -0.99), а корпоративная налоговая ставка имеет отрицательную асимметрию (-0.53), что указывает на наличие стран с низкими налоговыми ставками. Венчурные инвестиции и общий объем стартапов имеют очень высокую положительную асимметрию (6.33 и 6.88) и эксцесс (43.11 и 48.84), что подчеркивает наличие стран с экстремально высокими значениями этих показателей (например, США).

### Общие выводы для проекта:

1. ВВП на душу населения и средняя зарплата являются важными факторами для стартапов, так как они указывают на уровень экономического развития и покупательную способность населения.
2. Высокий уровень безработицы и инфляции могут быть негативными факторами, но в некоторых случаях они могут быть связаны с другими экономическими условиями, которые могут быть благоприятными для стартапов.
3. Индекс сложности ведения бизнеса и корпоративная налоговая ставка являются ключевыми факторами, влияющими на привлекательность страны для стартапов.
4. Венчурные инвестиции и общий объем стартапов являются важными показателями успешности стартап-экосистемы в стране.
5. **Анализ выбросов с помощью boxplot**

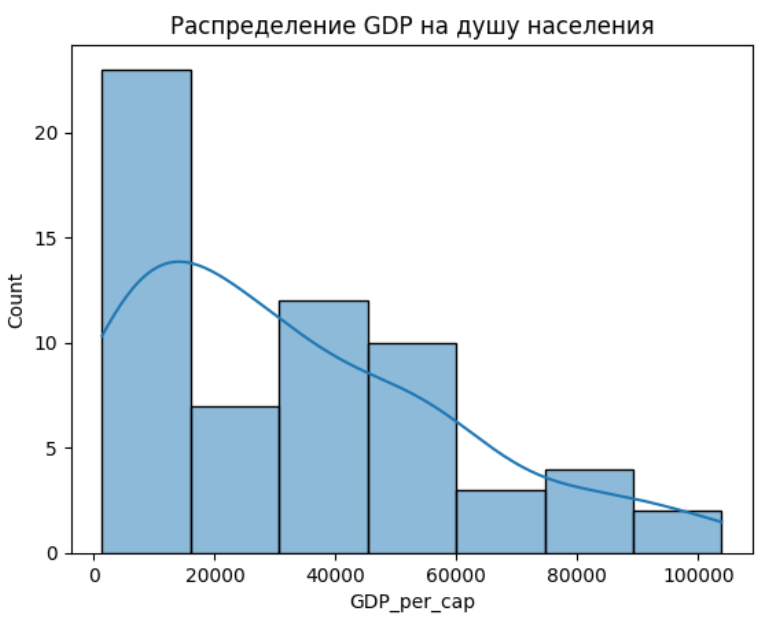


Boxplot выявил выбросы по нескольким ключевым показателям: высокий ВВП на душу населения (например, США, Швейцария), экстремально высокий уровень безработицы (например, Южная Африка), очень высокая инфляция (например, Венесуэла) и значительные венчурные инвестиции (например, США, Китай). Эти выбросы указывают на страны с уникальными экономическими условиями, которые могут быть как благоприятными, так и неблагоприятными для стартапов. Визуализация также подтвердила равномерное распределение индекса сложности ведения бизнеса и корпоративной налоговой ставки, что делает их важными факторами для анализа.

### **Выводы для проекта:**

### Страны с высоким ВВП на душу населения, высокой средней зарплатой и высокими венчурными инвестициями (например, США, Швейцария) могут быть наиболее привлекательными для стартапов.

**3. Визуализация распределения GDP на душу населения**

****

Гистограмма показывает распределение значений ВВП на душу населения по странам. KDE (Kernel Density Estimation) — это плавная кривая, которая аппроксимирует плотность распределения данных.

### **Интерпретация результатов визуализации**

Распределение ВВП на душу населения имеет положительную асимметрию (длинный правый хвост). Это означает, что большинство стран имеют относительно низкий ВВП на душу населения, но есть несколько стран с очень высоким ВВП.

Пик распределения находится в диапазоне 20,000–40,000, что указывает на то, что большинство стран имеют ВВП на душу населения в этом диапазоне.

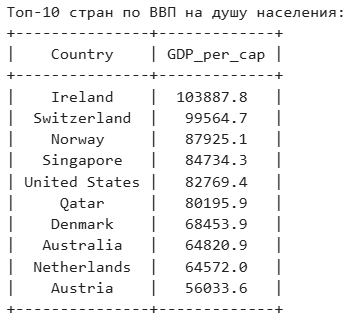
На графике виден длинный хвост вправо, что подтверждает наличие стран с экстремально высоким ВВП (например, США, Швейцария, Люксембург).

### Выводы для проекта:

1. Страны с высоким ВВП на душу населения (например, выше 60,000) могут быть наиболее привлекательными для стартапов, так как они обычно имеют развитую экономику, высокий уровень жизни и сильную инфраструктуру.
2. Выбросы (страны с очень высоким ВВП) требуют отдельного внимания, так как они могут сильно влиять на результаты анализа.

**4. Выявление топов 10 стран по различным показателям**

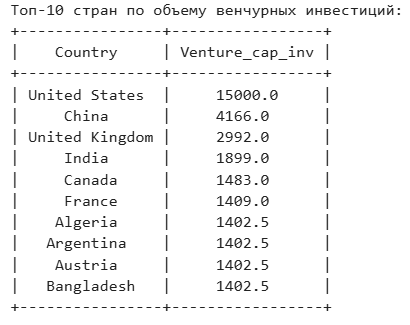
#### **Топ-10 стран по ВВП на душу населения**



ВВП на душу населения отражает общий уровень экономического развития страны, что напрямую влияет на возможности для стартапов.

Страны с высоким ВВП на душу населения (например, Ирландия, Швейцария, Норвегия) могут предоставить стартапам доступ к более богатым рынкам, квалифицированной рабочей силе и лучшей инфраструктуре. Это делает их привлекательными для запуска и развития бизнеса.

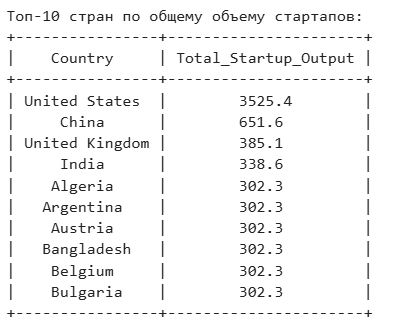
#### **Топ-10 стран по объему венчурных инвестиций**

****

Венчурные инвестиции — это ключевой источник финансирования для стартапов, и их объем напрямую влияет на успешность новых бизнесов.

Страны с высоким объемом венчурных инвестиций (например, США, Китай, Великобритания) предоставляют стартапам доступ к капиталу, что критически важно для их роста и развития.

#### **Топ-10 стран по общему объему стартапов**



Этот показатель отражает активность предпринимательской среды и показывает, насколько страна поддерживает инновации и новые бизнесы.

Страны с высоким объемом стартапов (например, США, Китай, Великобритания) имеют развитую экосистему для поддержки новых бизнесов, включая доступ к ресурсам, сетям и рынкам.

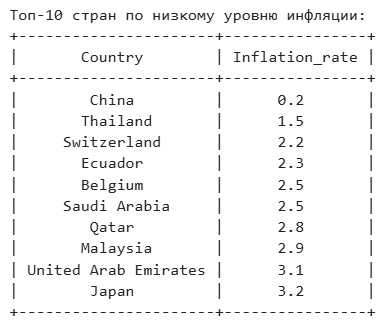
#### **Топ-10 стран по низкому уровню безработицы**



Уровень безработицы влияет на доступность рабочей силы и стабильность экономики, что важно для стартапов.

Страны с низким уровнем безработицы (например, Таиланд, ОАЭ, Япония) могут предложить стартапам доступ к квалифицированной рабочей силе, что критически важно для их успеха.

#### **Топ-10 стран по низкому уровню инфляции**



Инфляция влияет на стоимость ресурсов и стабильность экономики, что важно для стартапов, особенно на ранних этапах развития.

Страны с низким уровнем инфляции (например, Китай, Швейцария, Таиланд) предоставляют стартапам стабильную экономическую среду, что снижает риски и способствует долгосрочному планированию.

#### **Репрезентативный график с выбросами для 60 стран**

На данном графике мы видим выраженные выбросы в переменных GDP\_per\_cap, Inflation\_rate, Venture\_cap\_inv и Total\_Startup\_Output.

Несколько стран демонстрируют ВВП на душу населения выше $80,000, у других инфляция превышает 300%. Объем венчурных инвестиций в отдельных странах превышает 15,000 единиц. Объем стартапов в ряде стран также значительно выше средней выборки.

Присутствие этих выбросов искажает общую картину и мешает выявить закономерности между переменными. Поэтому мы переходим к следующему этапу — анализу без выбросов.

#### 

#### **Репрезентативный график без выбросов для 60 стран**

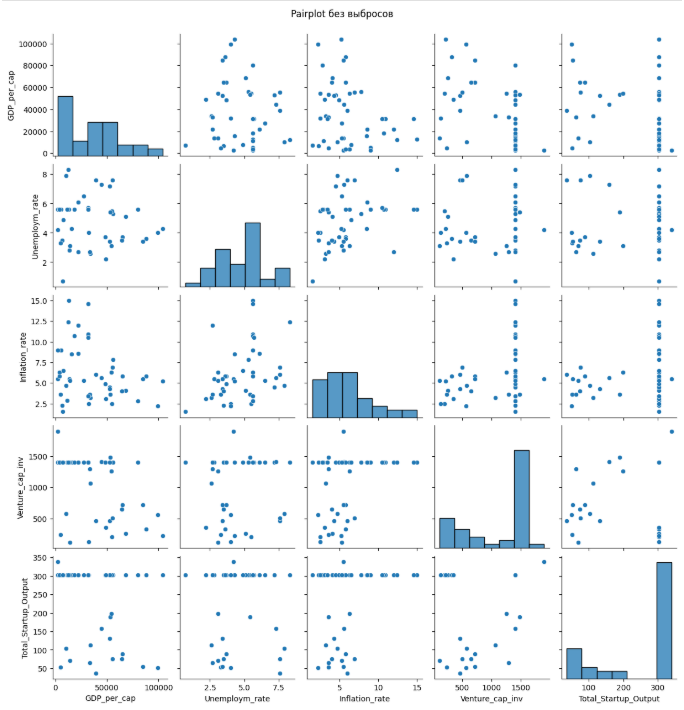
На этом графике — данные после удаления выбросов. Распределения становятся более равномерными и плотность скопления точек стала больше. Inflation\_rate варьируется в пределах от 2% до 15% без резких изменений.

Unemploym\_rate сглажен, большая часть стран находится в диапазоне от 3% до 8%.

Venture\_cap\_inv и Total\_Startup\_Output выглядят более реалистично, так как средние значения хорошо отражают ситуацию по большинству стран.

Присутствуют слабые положительные зависимости между Venture\_cap\_inv и Startup\_Output, а также между GDP\_per\_cap и Venture\_cap\_inv. Это имеет смысл, так как страны с высоким ВВП чаще инвестируют и создают больше стартапов.

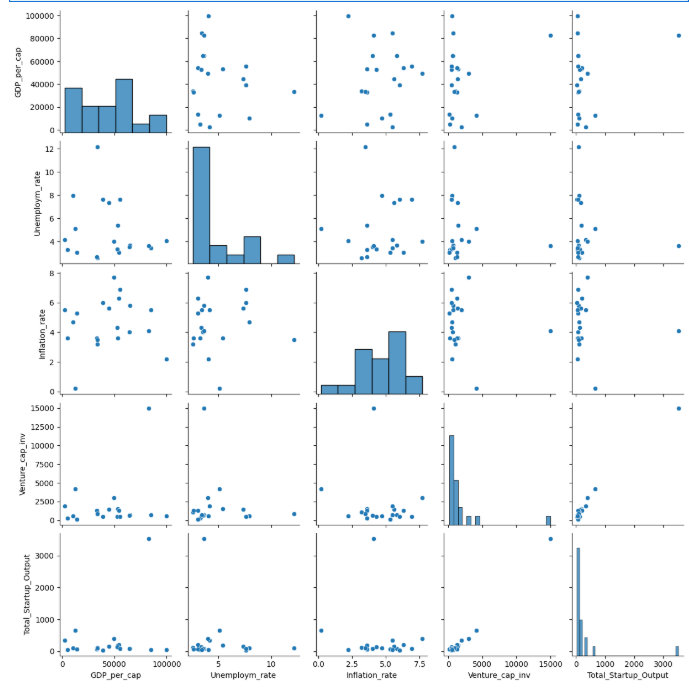
Удаление выбросов позволило более точно отразить взаимосвязи между переменными и подготовить данные к моделированию.



1. **Репрезентативный график для датасета с 20 странами**

На данном графике мы анализируем выборку из 20 стран, отобранных по критерию наличия данных по всем столбцам без заполнения пустых значений средним по столбцу.

Распределения становятся более сфокусированными — видно, что уровень инфляции и безработицы у этих стран в основном низкий или средний. ВВП на душу населения варьируется, но большая часть стран находится в диапазоне от $30,000 до $60,000. Venture\_cap\_inv и Startup\_Output находятся в ожидаемом диапазоне, и их взаимосвязь более выражена.



1. **Регрессионный анализ**

Общая гипотеза:

Экономические и бизнес-факторы страны значимо влияют на объем стартапов (Total\_Startup\_Output).

Конкретные гипотезы для каждой переменной:

**Венчурные инвестиции (Venture\_cap\_inv)**

H₀: Объем венчурных инвестиций не влияет на количество стартапов.

H₁: Увеличение венчурных инвестиций приводит к росту числа стартапов.

**Корпоративные налоги (Corporate\_Tax\_Rate)**

H₀: Уровень корпоративных налогов не влияет на объем стартапов.

H₁: Высокие налоги снижают количество стартапов.

**ВВП на душу населения (GDP\_per\_cap)**

H₀: ВВП не связан с объемом стартапов.

H₁: Страны с высоким ВВП имеют больше стартапов.

**Уровень безработицы (Unemploym\_rate)**

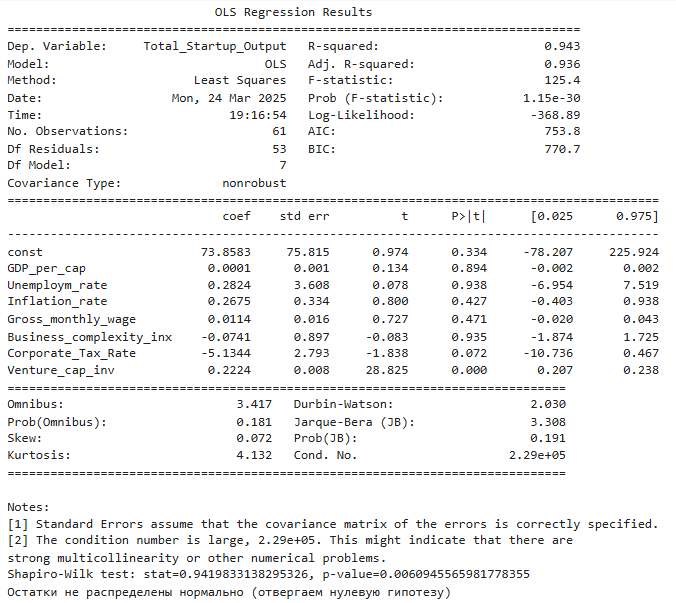
H₀: Безработица не влияет на стартапы.

H₁: Высокая безработица снижает количество стартапов.

**Инфляция (Inflation\_rate)**

H₀: Инфляция не влияет на стартапы.

H₁: Высокая инфляция подавляет стартапы.



#### **Проверка значимости переменных**

P-value для каждой переменной показывает, значимо ли она влияет на Total\_Startup\_Output (если p < 0.05).

Результаты:

1. **Венчурные инвестиции (Venture\_cap\_inv)**

Коэффициент: +0.2224

P-value: 0.000

Вывод: Объем венчурных инвестиций оказывает сильное положительное влияние на количество стартапов. Каждые дополнительные $1 млн инвестиций связаны с увеличением числа стартапов в стране на 0.22 единицы.

1. **Корпоративные налоги (Corporate\_Tax\_Rate)**

Коэффициент: -5.1344

P-value: 0.072

Вывод: Наблюдается пограничная статистическая значимость. Результаты позволяют предположить, что повышение корпоративных налогов на 1% может снижать количество стартапов примерно на 5.1 единицы. Однако из-за p-value, близкого к порогу значимости (0.073), этот вывод требует дополнительной проверки.

1. **Прочие факторы**

ВВП на душу населения, уровень безработицы, инфляция, средняя зарплата и индекс сложности ведения бизнеса показали статистически незначимые результаты.

Вывод: Эти переменные не оказывают существенного влияния на количество стартапов в рамках текущей модели.

#### Проверка качества модели

R² = 0.943: Модель объясняет 94.3% дисперсии данных — отличная точность.

F-statistic (p = 1.15e-30): Модель в целом значима.

#### **Итоговые выводы для проекта**

#### Главный фактор успеха:

#### Венчурные инвестиции (Venture\_cap\_inv) — единственный сильно значимый предиктор. Страны с высоким уровнем венчурных инвестиций (США, Китай, Великобритания) являются наиболее благоприятными для стартапов.

#### Второстепенные факторы:

#### Корпоративные налоги (Corporate\_Tax\_Rate) имеют пограничную значимость. Хотя влияние налогов не столь однозначно, страны с низкими корпоративными налогами могут иметь преимущество.

#### 

#### **Незначимые факторы:**

#### ВВП, безработица, инфляция, зарплата и сложность бизнеса не влияют на стартапы.

# **Ранжирование стран по составному индексу**

Мы применили два разных подхода для построения составного индекса Business Friendliness Index, на основе которого проводилось ранжирование стран по степени их привлекательности для стартапов.

Метод 1. Вычисления среднего арифметического нормализованных значений всех показателей

На начальном этапе анализа нами были выбраны следующие экономические показатели:

GDP\_per\_cap (ВВП на душу населения)

Unemploym\_rate (уровень безработицы)

Inflation\_rate (инфляция)

Gross\_monthly\_wage (средняя заработная плата)

Corporate\_Tax\_Rate (ставка корпоративного налога)

Venture\_cap\_inv (объём венчурных инвестиций)

Total\_Startup\_Output (общее количество стартапов)

Некоторые из показателей были обработаны для более корректной интерпретации.

Чтобы данные можно было сравнивать между собой, мы провели их нормализацию с помощью метода Min-Max Scaling и все показатели стали в диапазоне от 0 до 1.

Затем мы рассчитали индекс привлекательности бизнеса (Business Friendliness Index) для каждой страны путем вычисления среднего арифметического нормализованных значений всех показателей

Метод 2. Метод главных компонент (PCA)

Второй метод основан на анализе главных компонент (PCA — Principal Component Analysis). Это более продвинутый статистический метод, который позволяет:

Упростить структуру данных (уменьшить размерность)

Выделить направление, по которому различия между странами объясняются наибольшей дисперсией

Сначала мы также применили Min-Max нормализацию всех признаков. Затем на нормализованные данные был применён PCA с одной главной компонентой. Эта компонента представляла собой взвешенное сочетание всех исходных признаков, где веса подбирались автоматически на основе дисперсий в данных. То есть, признаки с большей вариативностью между странами получили больший вес.

Полученное значение главной компоненты и стало новым индексом Business Friendliness Index (PCA). Мы использовали его для альтернативного ранжирования стран и сравнения с первым методом.

# **Методология построения модели**

Мы также применили алгоритм машинного обучения Random Forest в рамках нашего проекта.

Мы ввели дополнительные веса для экономических факторов, которые были подобраны исходя из нашего теоретического анализа.

weights = {

'GDP\_per\_cap': 0.3,

'Unemploym\_rate': -0.2,

'Inflation\_rate': -0.1,

'Gross\_monthly\_wage': 0.1,

'Corporate\_Tax\_Rate': -0.1,

'Venture\_cap\_inv': 0.1,

'Total\_Startup\_Output': 0.1

}

Положительные веса означают, что увеличение значения этих показателей улучшает привлекательность страны для стартапов (например, высокий ВВП на душу населения).

Отрицательные веса говорят о том, что чем ниже показатель (например, инфляция или безработица), тем более благоприятна страна.

#### Принцип работы Random Forest в рамках проекта:

Этап обучения модели:

Random Forest представляет собой ансамбль деревьев решений, каждое из которых обучается на случайно отобранной подвыборке данных и случайном подмножестве признаков.

Например, отдельные деревья могут выделять наиболее важные признаки по-разному: одно дерево начинает разбиение по уровню ВВП на душу населения, другое — по инфляции, третье — по объему венчурных инвестиций. То есть, каждое дерево вносит уникальный вклад в общую модель.

После завершения обучения модель использует подход голосования: каждое дерево выносит свой прогноз (страна является или не является благоприятной для бизнеса — 1 или 0). Окончательный прогноз модели определяется большинством голосов, полученных от всех деревьев.

### **Проверка гипотезы**

На данной модели мы проверили нашу гипотезу “Страны с более высоким ВВП на душу населения более благоприятны для стартапов”.

Исходя из изначально подобранных весов, при проверке значимости показателей, ВВП оказался на первом месте. Мы решили проверить устойчивость модели и изменили веса, чтобы посмотреть как это влияет на значимость факторов и выводы. Несмотря на различные комбинации весов, показатель GDP\_per\_cap оставался фактором с самой высокой значимостью.

Это подтверждает нашу гипотезу, что высокий ВВП на душу населения создает более благоприятные для стартапов условия.

# **Результаты модели и индекса Business Friendliness Index**

Анализ составного индекса Business Friendliness Index

На основе расчета индекса были выявлены страны-лидеры, которые демонстрируют наибольший потенциал для развития стартапов. Среди них выделяются: США, Германия, Сингапур, Швейцария, Люксембург и Нидерланды с высоким уровнем ВВП на душу населения, развитой инфраструктурой и доступом к венчурному капиталу.

Эти страны характеризуются низкими уровнями инфляции и безработицы, а также благоприятной налоговой средой. Это делает их привлекательными для стартапа.

Результаты модели Random Forest

Использование Random Forest позволило получить четкое понимание того, как различные экономические факторы и их комбинации влияют на привлекательность страны для стартапов.

Результаты модели

Наиболее значимыми признаками, влияющими на успешность стран для развития стартапов, оказались:

* GDP\_per\_cap
* Gross\_monthly\_wage
* Unemploym\_rate

Важность этих признаков была значительно выше, чем у других факторов.

Модель Random Forest позволила нам при загрузке экономических показателей страны выявить, подходит ли она для стартапа (1 - подходит, 0 - не подходит).

# 

# **Выводы**

1. Один из главных факторов успеха стартапов – венчурные инвестиции

Фактор, который оказывает статистически значимое положительное влияние на количество стартапов в стране, — объем венчурных инвестиций (Venture\_cap\_inv), согласно результатам регрессионного анализа (p-value < 0.001). Страны с высоким уровнем венчурного финансирования, такие как США и Китай имеют наибольший потенциал для успешного запуска стартапа.

2. Высокий уровень экономического развития (ВВП на душу населения)

Проведенная нами проверка гипотезы показала, что GDP\_per\_cap (ВВП на душу населения), всегда был самым важным фактором, независимо от различных комбинаций веcов. Согласно визуальному и разведывательному анализу, страны с высоким ВВП, такие как Швейцария, Германия и США, являются более привлекательными для бизнеса.

3. Роль стабильной экономической среды и налогового режима

Регрессионный анализ показал, что корпоративная налоговая ставка имеет пограничную значимость (Corporate\_Tax\_Rate, p-value = 0,073). Несмотря на то, что статистические данные не полностью достоверны, результаты дают нам предположить, что умеренные или низкие корпоративные налоги положительно влияют на привлекательность страны для стартапов.

Самые привлекательные страны (Нидерланды, Германия и Швейцария) демонстрируют стабильность экономических условий (низкая инфляция, умеренные налоги).

4. Негативные факторы для стартапов

Согласно нашему исследованию, страны с очень высокой инфляцией и нестабильной экономикой, такие как Венесуэла и Южная Африка, значительно менее привлекательны для бизнеса. Высокие риски и нестабильная экономика мешают бизнесу.

## **Общий итог исследования**

Согласно проведенному нами тщательному анализу, США, Германия, Сингапур, Швейцария, Люксембург и Нидерланды являются наиболее привлекательными странами для создания стартапов, поскольку они имеют наилучший набор характеристик:

1. Высокого уровня венчурных инвестиций
2. Высокого ВВП на душу населения и развитой инфраструктуры
3. Стабильной экономической среды
4. Умеренной налоговой политики.

#### 

#### 