Тема моей курсовой работы аналитика турпотоков в России с определение факторов влияния. Актуальность данной работы состоит в том, что анализ данных играет все более важную роль в различных сферах человеческой деятельности, включая туризм. С развитием информационных технологий и доступностью больших объемов данных, становится возможным проводить глубокий и комплексный анализ туристических потоков и факторов, влияющих на развитие туристической индустрии.

--------------------------------------------------------------------------------------------

В связи с этим целью данной работы является проведение анализа турпотоков в России за определенный период времени с помощью языка R, а также изучение туриндустрии с целью выявления основных тенденций, паттернов и факторов, влияющих на развитие отрасли.

Для реализации поставленной цели предполагается решить следующие задачи.

1. Рассмотреть методы интеллектуального анализа данных;
2. Изучение сферы исследования – турпотоки в России;
3. Проанализировать турпотоки и сделать выводы;
4. Проанализировать туристическую индустрию и сделать выводы.

--------------------------------------------------------------------------------------------

Разберем сначала, что такое **Data Mining**. Технология **Data Mining** заключается в поиске неподготовленных, безызвестных ранее данных, а также полезных и доступных к использованию знаний для принятия решений. Так как данная технология предназначена для функционирования в глобальных базах данных, то на практике она имеет несколько закономерностей, а именно: **неочевидных, объективных и практически полезных.**

**Неочевидные закономерности** подразумевают под собой закономерности, которые нельзя обнаружить стандартным путем анализа.

**Объективные** — данные закономерности полностью соответствую действительности, чего нельзя сказать об экспертном мнении, которое является субъективным.

**Практически** полезные закономерности говорят нам о том, что конкретным значениям можно найти прямое практическое применение.

--------------------------------------------------------------------------------------------

Далее рассмотрим, что такое данные и какие же нужны требования к их обработке. **Данные** — это факты, тексты, графики числовые выражения, документы, и другого рода сегменты. Данные должны быть конвертируемыми, то есть, способны менять свою форму для необходимого типа хранения. Можно сказать, данные — набор необработанной информации. Все это может быть получено в результате практической деятельности.

Существует и определенная специфика требований к обработке данных:

1. Неограниченный объем.
2. Разнородность.
3. Корректные результаты.
4. Простые инструменты для обработки.

--------------------------------------------------------------------------------------------

Традиционный процесс сбора данных состоит из следующих пунктов:

1. Анализ предметной области.
2. Постановка задачи.
3. Подготовка данных.
4. Построение моделей.
5. Проверка и оценка.
6. Выбор.
7. Применение.
8. Коррекция и обновление.

--------------------------------------------------------------------------------------------

**Качество данных** — критерий, определяющий полноту и точность предоставляемых данных, а также, возможность их интерпретации. **Данные** **высокого качества** — это полные и точные данные, обеспечивающие исследователю на их фоне составить полномерную корректно функционирующую модель способную к интерпретации и принятию решений. **Данные низкого качества**, соответственно, таких характеристик не имеют, включая неверный формат.

Наиболее распространенные виды так называемых «грязных» данных:

1. Пропущенные значения.
2. Дубликаты.
3. Шумы и выбросы.

--------------------------------------------------------------------------------------------

Теперь перейдем непосредственно к анализируемой сфере исследования – туристические потоки в России. **Турпотоки в России** представляют собой потоки туристов, как въезжающих в страну, так и выезжающих из нее.

Анализ турпотоков в России является важным исследовательским направлением, поскольку позволяет понять и оценить динамику и характер туристического потока в разных регионах страны. Это помогает выявить популярные направления и места, наиболее востребованные виды туризма, а также понять предпочтения и поведение туристов.

--------------------------------------------------------------------------------------------

**Виды турпотоков** включают различные направления и характеристики движения туристов. В контексте анализа турпотоков в России выделим несколько основных видов:

1. Въездные турпотоки: Это потоки туристов, приезжающих в Россию из других стран. Въездные турпотоки могут быть связаны с различными целями путешествия, включая туризм, деловые визиты, культурный обмен и другие.
2. Выездные турпотоки: Это потоки туристов, покидающих Россию и отправляющихся в другие страны. Выездные турпотоки также отражают предпочтения и потребности российских туристов.
3. Внутренние турпотоки: Это потоки туристов, путешествующих внутри России.

**Туристическая индустрия** включает в себя разнообразные организации и предприятия, которые осуществляют деятельность в сфере туризма.

1. Прибыльные и убыточные организации.
2. Выручка: Включает доходы от проживания, питания, транспортировки, экскурсий и других туристических активностей.
3. Прибыль и убыток: Прибыль является разницей между выручкой и затратами, показывая финансовый успех организации. Убыток, напротив, свидетельствует о негативных финансовых результатах.
4. НФР (Сальдированный финансовый результат): НФР включает в себя инвестиции в развитие и поддержку туристической инфраструктуры, включая строительство, ремонт и обновление объектов туризма. Прибыль минус убыток.
5. Ввод в действие объектов туризма: Включает оценку строительства и открытия новых отелей, гостиниц, курортных комплексов, музеев, парков и других объектов, которые привлекают туристов.

--------------------------------------------------------------------------------------------

Перейдем непосредственно к анализу. На слайде вы можете увидеть график, показывающий количество въезжающих граждан из каждой страны в период с 2014-2022 года. Можно выделить 3 страны лидера. **Украина** привлекает внимание своей географической близостью, культурными и историческими связями с Россией, а также благодаря разнообразным туристическим предложениям. К сожалению, в связи с последними событиями, в 2022 году количество въезжающих с Украины серьезно сократилось. Благоприятные экономические и культурные отношения между **Казахстаном и Россией** способствуют активным туристическим пересечениям, включая посещение достопримечательностей, бизнес-туризм и шопинг. Благодаря развитию туристической индустрии и улучшению транспортных связей с Россией, жители **Узбекистана** становятся все более частыми гостями в страну. Конечно же, не малую роль играет жители, которые приезжают в Россию на заработки и, в большинстве случае, остаются на постоянное место жительство.

--------------------------------------------------------------------------------------------

Граждане Финляндии, Польши и Испании являются частыми посетителями России и активно совершают выездные поездки в нашу страну. Это можно объяснить близостью к границам России и интересу к местной культуре.

--------------------------------------------------------------------------------------------

Рассмотрит график выездных турпоездок. Одним из ведущих направлений, которое привлекает выезжающих российских туристов, является также **Украина**. Еще одним интересным трендом на графике является растущая популярность **Турции** среди российских туристов начиная с 2018 года. Турция привлекает своим прекрасным климатом, привлекательными курортами, богатой историей и гостеприимством своих жителей. Заметной тенденцией на графике является также выездные поездки в **Абхазию**, особенно с 2018 года. Россияне оценивают близость курортов Абхазии к Сочи, что упрощает переход в эту страну.

--------------------------------------------------------------------------------------------

На графике выездных турпотоков также наблюдается популярность выезда граждан в **Финляндию**. Это обусловлено тем, что Финляндия граничит с Санкт-Петербургом, что также упрощает въезд в страну.

--------------------------------------------------------------------------------------------

Далее посмотрим на внутренние потоки за 2022 год. Анализируя представленный график, можно выделить несколько регионов, которые являются особенно популярными среди путешественников внутри России. **Московская область, в частности город Москва,** притягивает огромное количество туристов. Богатая история, культурное наследие, величественные архитектурные сооружения и динамичная атмосфера делают Москву очень популярным туристическим направлением. **Ленинградская область, с прекрасным городом Санкт-Петербургом во главе**, также привлекает внимание множества путешественников благодаря историческим достопримечательностям, великолепной архитектуре и уникальной атмосфере. **Краснодарский край**, с его прекрасными курортами и природными ландшафтами, также является магнитом для туристов. Солнечные пляжи Черного моря, уникальные горные пейзажи Кавказа и гостеприимство местных жителей делают этот регион идеальным местом для отдыха и путешествий.

--------------------------------------------------------------------------------------------

Далее проведем кластеризацию данных о путешествиях внутри России. В ходе анализа алгоритм выделил три кластера.

1) Наиболее посещаемые области.

2) Не очень популярные, но все равно привлекающие внимание путешественников.

3) Редко посещаемые области.

Как было сказано ранее, наиболее часто посещаемыми области являются **Московская область, Ленинградская область и Краснодарский край**. Во вторую группу входят, такие области как **Владимирская область, Ставропольский край, Тюменская область** и многие другие. Северные области, такие как **Алтайский край, Новосибирская область и Красноярский край**, выделяются как менее популярные направления для туристов в России.

--------------------------------------------------------------------------------------------

Для более детального анализа путешествий по округам, можно разделить год на четыре сезона, с учетом климатических условий. **Четвертый и первый** кварталы, включающие октябрь, ноябрь, декабрь, январь, февраль, март, можно отнести к холодным сезонам, в то время **второй и третий** - апрель, май, июнь, июль, август, сентябрь - можно отнести к теплым сезонам.

Лидирующие позиции в теплые сезоны также занимают **Краснодарский край, Ленинградская и Московская области**. В холодное время года **Московская и Ленинградская** области продолжают привлекать туристов, уступая лишь небольшую долю популярности в сравнении с теплыми сезонами. **Краснодарский край** также остается привлекательным благодаря своим горнолыжным курортам. В то же время, **другие** регионы не привлекают такого большого числа посетителей как в теплое, так и в холодное время года, в частности северные регионы не являются желанными точками остановки для путешественников.

--------------------------------------------------------------------------------------------

Перейдем к анализу Туриндустрии. На графике мы можем увидеть число организаций за последние 9 лет.

--------------------------------------------------------------------------------------------

На следующих двух число прибыльных и убыточных организаций. При анализе выше описанных графиков становится очевидным влияние пандемии Covid-19 на туристическую отрасль в стране. В 2020 году отмечается резкое снижение числа туристов и путешественников, что является результатом ограничений в стране.

--------------------------------------------------------------------------------------------

На следующих четырех графиках мы наблюдаем выручку, прибыль, убыток и НФР (прибыль минус убыток). Мы можем наблюдать серьезные убытки и снижение выручки в 2022 году, вызванные пандемией. Однако, также стоит отметить интересную тенденцию в 2015 году, когда отмечаются колоссальные убытки со стороны туристических организаций. Вероятно, это связано с серией терактов, произошедших в разных частях мира, что вызвало опасения и неуверенность у путешественников.

--------------------------------------------------------------------------------------------

На графике видим ввод в действие объектов туризма. За промежуток девяти лет наблюдается стабильное открытие гостиниц и отелей в значительном количестве, указывающее на широкую доступность размещения для туристов. Однако, кемпинги и мотели остаются наименее популярными вариантами размещения. В нашей стране данные виды объектов туризма не привлекают массового числа отдыхающих. Санатории и туристические базы, в свою очередь, удерживают свою позицию каждый год.

--------------------------------------------------------------------------------------------

В результате работы были изучены

* турпотоки в России, включая въездные и выездные турпоездки, а также путешествия внутри страны,
* ведущие роли различных стран в числе приезжих и выезжающих туристов,
* популярные туристические регионы,
* тенденции упадка в туристической индустрии в связи с различными факторами,
* углубленно изучен язык для анализа данных R, его методы построения графиков и алгоритмы кластеризации данных.