**Выступление по теме "Анализ экологии воздуха Курского края"**

Добрый день, уважаемые коллеги и гости.

Меня зовут Вячеслав Гамов, а моего коллегу - Владислав Фёдоров. Мы представляем результаты нашего исследования на тему "Анализ экологии воздуха Курского края", выполненного в рамках программы Samsung IT Academy.

**01 Введение (Жирным** – **слайд**, *Курсив - подтема***)**

В данном исследовании мы провели детальный анализ качества воздуха в Курском крае, выявили основные экологические проблемы, связанные с загрязнением атмосферы, и представили прогнозы по ожидаемой экологической ситуации в регионе.

**02 Актуальность проекта**

В настоящее время проекты, связанные с анализом экологического состояния воздуха, становятся все более актуальными в свете нарастающих проблем окружающей среды.

Понимание проблем загрязнения и необходимость улучшения экологии подчеркивают важность доступа к точной и объективной информации, которую можно получить с помощью Big Data. Эти технологии позволяют собирать и анализировать огромные объемы данных, что способствует выявлению скрытых закономерностей и тенденций.

**03 Цели и задачи**

*Цель проекта*

Целью проекта является анализ больших данных о состоянии экологии воздуха в Курской области и составление наглядной статистики с графиками и диаграммами. Эти данные помогут выявить основные источники загрязнения, а также определить временные и пространственные тенденции в изменении качества воздуха.

*Задачи проекта*

* Сбор источников содержащих потенциально полезную информацию и наборов данных;
* Сбор данных о загрязнении воздуха;
* Анализ данных, составление статистики;
* Поиск корреляций, оформление статистики;
* Визуализация результатов анализа и создание отчетов;

**04 Целевая аудитория**

*Государственные органы и учреждения*

Государственные органы и учреждения - целевая аудитория проекта для разработки программ по снижению загрязнения воздуха, включая ограничение выбросов и поддержку экологически чистого транспорта.

*Научные и исследовательские институты*

Научные исследователи важны для анализа данных о загрязнении воздуха, выявлении тенденций и понимании воздействия загрязнений на окружающую среду и здоровье людей.

*Общественные организации и НПО*

НПО и общественные организации занимаются защитой окружающей среды и здоровья населения. Они могут использовать данные о качестве воздуха для лоббирования изменений в политике, нормативных актах и законах.

*Широкая общественность*

Информирование общественности о состоянии воздуха может помочь снизить риски заболеваний, таких как астма, аллергии и сердечно-сосудистые заболевания.

**05 Набор данных**

Для данного проекта использовались данные из архива недавних исследований качества воздуха в Курской области. Эти данные были проанализированы на официальном сайте Министерства природных ресурсов Курской области, а также с использованием онлайн-сервисов оценки качества воздуха. Кроме того, в ходе сотрудничества с Курским Гидрометцентром были получены сведения о загрязнении воздуха за период с 2002 по 2024 год.

**06 Используемые инструменты (бла-бла-бла…)**

**07 Инструменты для прогнозирования**

Прогнозирование было выполнено с помощью PySpark, библиотек PySpark.ML. Была применена модель линейной регрессии. Оценка точности моделей по критерию AUC-ROC, равна 0.72 Кроме того, было рассмотрено использование других алгоритмов машинного обучения, таких как деревья решений или градиентный бустинг, но их использование приводило к ухудшению алгоритма прогнозирования.

**08 Аммиак**

Средняя концентрация аммиака в воздухе: на протяжении рассматриваемого периода средняя концентрация аммиака в воздухе находилась в пределах от 0,00442 мг/м³ (2021 год) до 0,02450 мг/м³ (2023 год). Средние значения по годам колебались, но все они оставались значительно ниже безопасного уровня в 0,04 мг/м³. Максимальная концентрация аммиака в воздухе: Максимальные зафиксированные концентрации аммиака также варьировались, с пиком в 0,06000 мг/м³ в 2002 году. Максимальные значения за весь период не превышали безопасную концентрацию в 0,04 мг/м³ только в одном году — в 2002 году.

Тренды и отклонения: В 2023 году наблюдается заметный рост средней концентрации аммиака до 0,02450 мг/м³, что является самым высоким значением за весь рассматриваемый период. Годы с наименьшими средними концентрациями аммиака — 2020 и 2021, что может быть связано с изменениями в хозяйственной деятельности в связи с пандемией COVID-19.

*Прогноз*

Согласно проведенному анализу, был составлен прогноз уровня аммиака в атмосфере на два года: В 2025 году прогнозируется увеличение количества частиц аммиака до 0.061 что будет являться критическим уровнем по стандарту ПДК Далее, в 2026 году ожидается снижение концентрации взвешенных частиц, по сравнению с прошлым годом, до уровня 0.05 мг/m3

**09 Формальдегид**

Средняя концентрация формальдегида в воздухе: В 2002 году: 0.01538 мг/м³ В 2023 году: 0.02700 мг/м³ На протяжении времени с 2002 по 2023 год средняя концентрация формальдегида в воздухе колебалась от 0.00600 до 0.02725 мг/м³. Самая высокая средняя концентрация была зафиксирована в 2022 и 2023 годах. Максимальная зафиксированная концентрация формальдегида в воздухе: Максимальные значения колебались значительно с более низкими значениями в начале периода и достигали более высоких значений в конце. Например, в 2016 году было зафиксировано значение 0.12550 мг/м³, что является значительно выше, чем в другие годы.

Выводы: Средняя концентрация формальдегида часто превышает безопасный уровень 0.01 мг/м³, особенно в более поздние годы (например, с 2015 по 2023 годы). В период с 2002 по 2023 годы наблюдается колебание средней и максимальной концентраций формальдегида в воздухе.

*Прогноз*  
Согласно проведенному анализу, был составлен прогноз уровня аммиака в атмосфере на два года: В 2025 году прогнозируется увеличение количества частиц формальдегида до 0.06 что будет являться критическим уровнем по стандарту ПДК Далее, в 2026 году ожидается снижение концентрации формальдегида, по сравнению с прошлым годом, до уровня 0.05 мг/m3

**10 Взвешенные вещества**

Максимальные значения концентрации взвешенных веществ: Максимальная концентрация взвешенных веществ в воздухе варьируется от 0,050417 мг/м³ в 2022 году до 0,115917 мг/м³ в 2017 году. Наибольшие значения были зафиксированы в 2016 и 2017 годах, что может указывать на периоды повышенного загрязнения воздуха.

Выводы: В целом, наблюдается колебание значений концентрации воздушных веществ с годами, что может быть обусловлено различными факторами, включая изменения в промышленности, транспорте, и климатические условия.

*Прогноз*

Согласно проведенному анализу, был составлен прогноз уровня взвешенных веществ в атмосфере на два года: В 2025 году прогнозируется уменьшение количества взвешенных частиц до 0.0624 по сравнению с 2024 годом. Далее, в 2026 году ожидается дальнейшее снижение концентрации взвешенных частиц, указывающее на общее уменьшение уровня пыли в городе.

**11 Сравнение загрязненности округов**

Наибольшая концентрация взвешенных веществ и формальдегида была обнаружена на участке №36, что свидетельствует о наивысшем уровне загрязнения Центрального района по сравнению с другими районами. Средний уровень загрязнения характеризует Железнодорожный район. Наименьшее количество загрязняющих частиц в воздухе зафиксировано в Сеймском районе, что делает его самым чистым районом Курска.

**Заключение**

В ходе нашего исследования были использованы технологии обработки и анализа данных, включая библиотеки PySpark и PySpark.ML. Была применена модель линейной регрессии для прогнозирования уровней загрязнителей воздуха в Курском крае. Использование этих инструментов позволило нам собрать и проанализировать большие объемы данных, выявив скрытые закономерности и тенденции. Анализ показал, что средние концентрации аммиака и формальдегида колеблются в зависимости от года, причем уровни формальдегида часто превышают безопасный уровень. Прогнозы на 2025 и 2026 годы указывают на критические уровни загрязнителей, что требует принятия мер для улучшения качества воздуха.

**Спасибо за внимание**