

ДЕПАРТАМЕНТ ОБРАЗОВАНИЯ И НАУКИ ГОРОДА МОСКВЫ

Государственное автономное образовательное учреждение

высшего образования города Москвы

«Московский городской педагогический университет»

(ГАОУ ВО МГПУ)

Институт цифрового образования

Департамент информатики, управления и технологий

Практическая(лабораторная) работа № 3.1

по дисциплине «Платформы Data Engineering»

Выполнил:

студент группы БД-251м

Направление подготовки/Специальность

38.04.05 - Бизнес-информатика

Войт Иван Иванович

(Ф.И.О.)

Проверил:

Доцент департамента информатики, управления и технологий, доктор

экономических наук

(ученая степень, звание)

Босенко Т.М.

(Ф.И.О.)

Москва 2025

Оглавление	
Введение	3
Задачи:.....	3
Описание выбранного варианта.....	3
Процесс разработки	4
1. Подключение к источнику данных	4
2. Создание и подготовка датасета	4
3. Разработка чартов.....	5
4. Настройка селекторов.....	6
Заключение.....	7

Введение

Цель работы: Разработать интерактивный аналитический дашборд для мониторинга и анализа эффективности цепочки поставок с использованием Yandex DataLens.

Задачи:

- Создать дашборд в Yandex DataLens на основе индивидуального варианта
- Реализовать ключевые метрики анализа цепочки поставок
- Применить принципы эффективного дизайна дашбордов
- Обеспечить интерактивность для глубокого анализа данных

Описание выбранного варианта

Тема: Анализ цепочки поставок.

Набор данных: bdt-mba-supply-chain-dataset

Основные метрики: Надежность поставщиков, оборачиваемость запасов, уровень сервиса.

Рекомендуемые визуализации: Матрица рисков, диаграмма Санкея движения товаров, карта поставщиков

Процесс разработки

1. Подключение к источнику данных

В качестве источника данных использовался CSV-файл `bdt_mba_supplychain_dataset_2024.csv`, содержащий информацию о 500 активах цепочки поставок с следующими полями:

Таблица 1 – Поля источника данных

Asset_ID	Идентификатор актива
Location	Местоположение
Temperature, Vibration	Технические параметры
Last_Maintenance	Дата последнего ТО
Condition_Score -	Оценка состояния
Resource_Utilization	Использование ресурсов
Delivery_Efficiency	Эффективность доставки
Downtime_Hours	Время простоя
Inventory_Level	Уровень запасов
Logistics_Cost	Логистические затраты

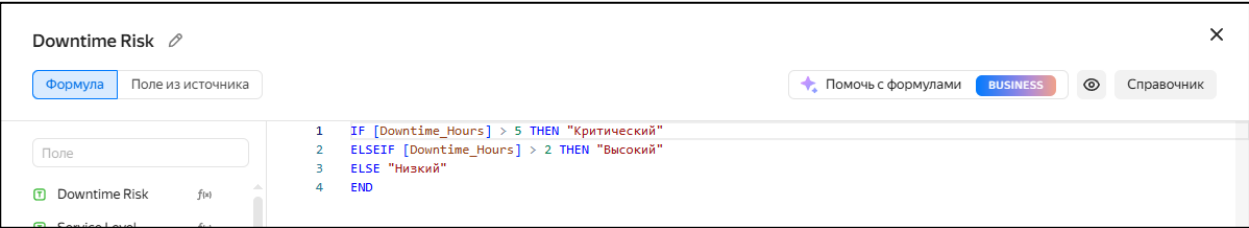
2. Создание и подготовка датасета

Все поля были проверены и приведены к корректным типам данных:

- Строковые поля: `Asset_ID`, `Location`, `Inventory_Level`
- Числовые поля: `Temperature`, `Vibration`, `Condition_Score` и др.
- Дата/время: `Last_Maintenance`, `Timestamp`

Далее были созданы следующие вычисляемые поля:

Риск простоя (Downtime Risk):



Надежность актива (Asset Reliability):

Asset Reliability

ФормулаПоле из источника

Помочь с формуламиBUSINESSСправочник

Поле

Downtime Risk

f(n)

```
1 IF [Condition_Score] > 80 THEN "Высокая"
2 ELSEIF [Condition_Score] > 60 THEN "Средняя"
3 ELSE "Низкая"
4 END
```

Уровень обслуживания (Service Level):

Service Level

ФормулаПоле из источника

Помочь с формуламиBUSINESSСправочник

Поле

Downtime Risk

f(n)

```
1 IF [Delivery_Efficiency] >= 90 THEN "Превосходный"
2 ELSEIF [Delivery_Efficiency] >= 80 THEN "Высокий"
3 ELSEIF [Delivery_Efficiency] >= 70 THEN "Средний"
4 ELSE "Низкий"
5 END
```

Эффективность логистики (Logistics Efficiency Ratio):

Logistics Efficiency Ratio

ФормулаПоле из источника

Помочь с формуламиBUSINESSСправочник

Поле

Downtime Risk

f(n)

```
1 [Delivery_Efficiency] / [Logistics_Cost]
```

3. Разработка чартов

Индикаторы ключевых метрик:

- Средняя эффективность доставки: `AVG ([Delivery_Efficiency])`
- Среднее время простоя: `AVG ([Downtime_Hours])`
- Общие логистические затраты: `SUM ([Logistics_Cost])`
- Общее количество активов: `COUNT ([Asset_ID])`

Матрица рисков:

Обоснование выбора - Классическая матрица рисков позволяет оценить активы по двум параметрам - вероятности сбоя (состояние) и воздействию (простой)

Настройка:

X: Condition_Score

Y: Downtime_Hours

Цвет: Риск простоя (Downtime Risk)

Размер точек: Logistics_Cost

Точки: Asset_ID

Анализ запасов и сервиса (Сводная таблица):

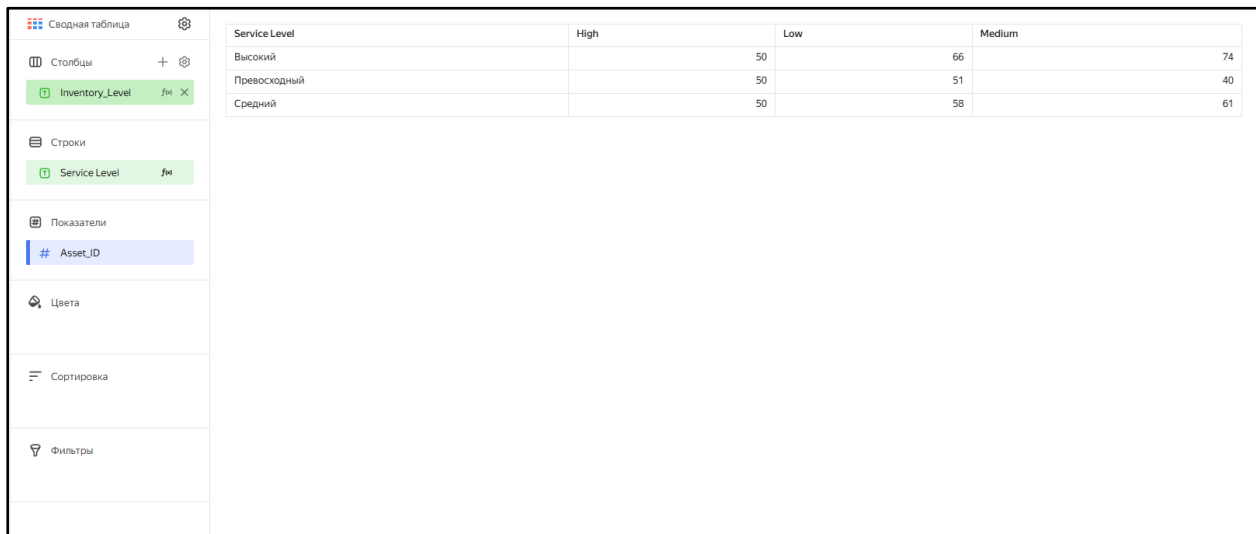
Обоснование выбора - показывает четкое распределение активов между уровнями запасов и сервиса

Настройка:

Строки: Уровень обслуживания (Service Level)

Столбцы: Inventory_Level

Показатель: COUNT ([Asset_ID])



Service Level	High	Low	Medium
Высокий	50	66	74
Превосходный	50	51	40
Средний	50	58	61

Детальный анализ по локациям (Сводная таблица):

Обоснование выбора - предоставляет детализированную информацию по всем локациям

Настройка:

Строки: Location

Показатели: COUNT ([Asset_ID]), AVG ([Condition_Score]),
AVG ([Downtime_Hours]), SUM ([Logistics_Cost])

4. Настройка селекторов

Для обеспечения интерактивности были добавлены селекторы:

- Уровень риска (Риск простоя)
- Уровень запасов (Inventory_Level)
- Надежность активов
- Уровень сервиса

Заключение

В ходе работы был успешно разработан интерактивный дашборд для анализа цепочки поставок в Yandex DataLens. Основные достижения:

- Реализованы ключевые метрики анализа цепочки поставок через вычисляемые поля
- Создана комплексная визуализация включающая матрицу рисков, анализ запасов и детализацию по локациям
- Обеспечена высокая интерактивность через систему связанных селекторов
- Применены принципы эффективного дизайна для создания понятного и информативного интерфейса

Дашборд позволяет бизнес-пользователям:

- Выявлять проблемные активы через матрицу рисков
- Анализировать зависимость между запасами и уровнем сервиса
- Детализировать анализ по конкретным локациям и категориям
- Принимать обоснованные решения по оптимизации цепочки поставок

Работа подтвердила, что Yandex DataLens является эффективным инструментом для быстрого создания операционных дашбордов, особенно для задач мониторинга и тактического анализа.

Ссылки

Дашборд в Yandex DataLens: [<https://datalens.ru/z9ib96ihe98wk>]

Git-репозиторий с работой: [[https://github.com/youngvoyt/DEP-MGPU/tree/ccc21610617262d270c1bcb984f4a608465f5077/Module03/Lab%203.](https://github.com/youngvoyt/DEP-MGPU/tree/ccc21610617262d270c1bcb984f4a608465f5077/Module03/Lab%203.1)

[1](#)]