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**](#_Toc212350392)

[**Задачи: 3**](#_Toc212350393)

[**Описание выбранного варианта 3**](#_Toc212350394)

[**Процесс разработки 4**](#_Toc212350395)

[**1. Подключение к источнику данных 4**](#_Toc212350396)

[**2. Создание и подготовка датасета 4**](#_Toc212350397)

[**3. Разработка чартов 5**](#_Toc212350398)

[**4. Настройка селекторов 6**](#_Toc212350399)

[**Заключение 7**](#_Toc212350400)

# **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**

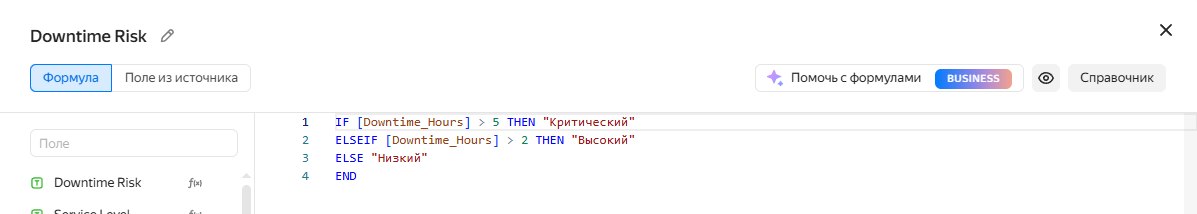
|  |  |
| --- | --- |
| Asset\_ID | Идентификатор актива |
| Location | Местоположение |
| Temperature, Vibration | Технические параметры |
| Last\_Maintenance | Дата последнего ТО |
| Condition\_Score - | Оценка состояния |
| Resource\_Utilization | Использование ресурсов |
| Delivery\_Efficiency | Эффективность доставки |
| Downtime\_Hours | Время простоя |
| Inventory\_Level | Уровень запасов |
| Logistics\_Cost | Логистические затраты |

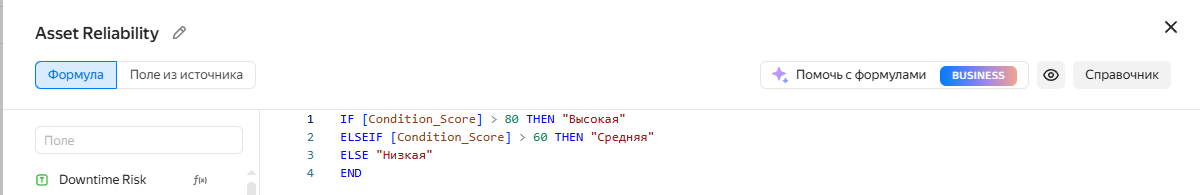
## **2. Создание и подготовка датасета**

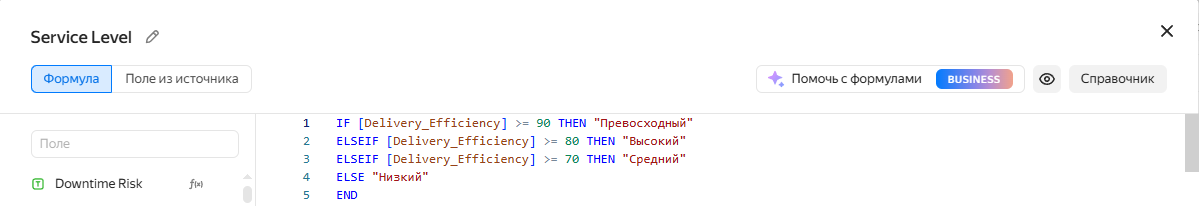
Все поля были проверены и приведены к корректным типам данных:

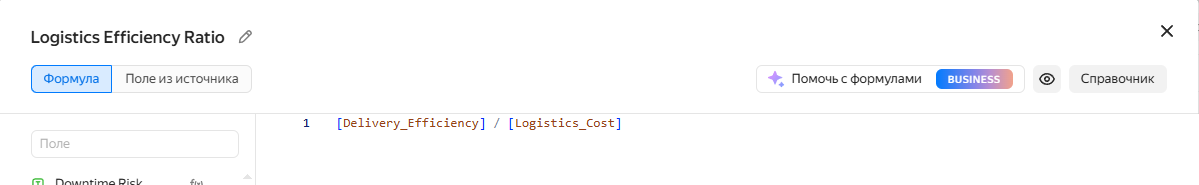
* Строковые поля: Asset\_ID, Location, Inventory\_Level
* Числовые поля: Temperature, Vibration, Condition\_Score и др.
* Дата/время: Last\_Maintenance, Timestamp

Далее были созданы следующие вычисляемые поля:

Риск простоя (Downtime Risk):

Надежность актива (Asset Reliability):

Уровень обслуживания (Service Level):

Эффективность логистики (Logistics Efficiency Ratio):

## **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

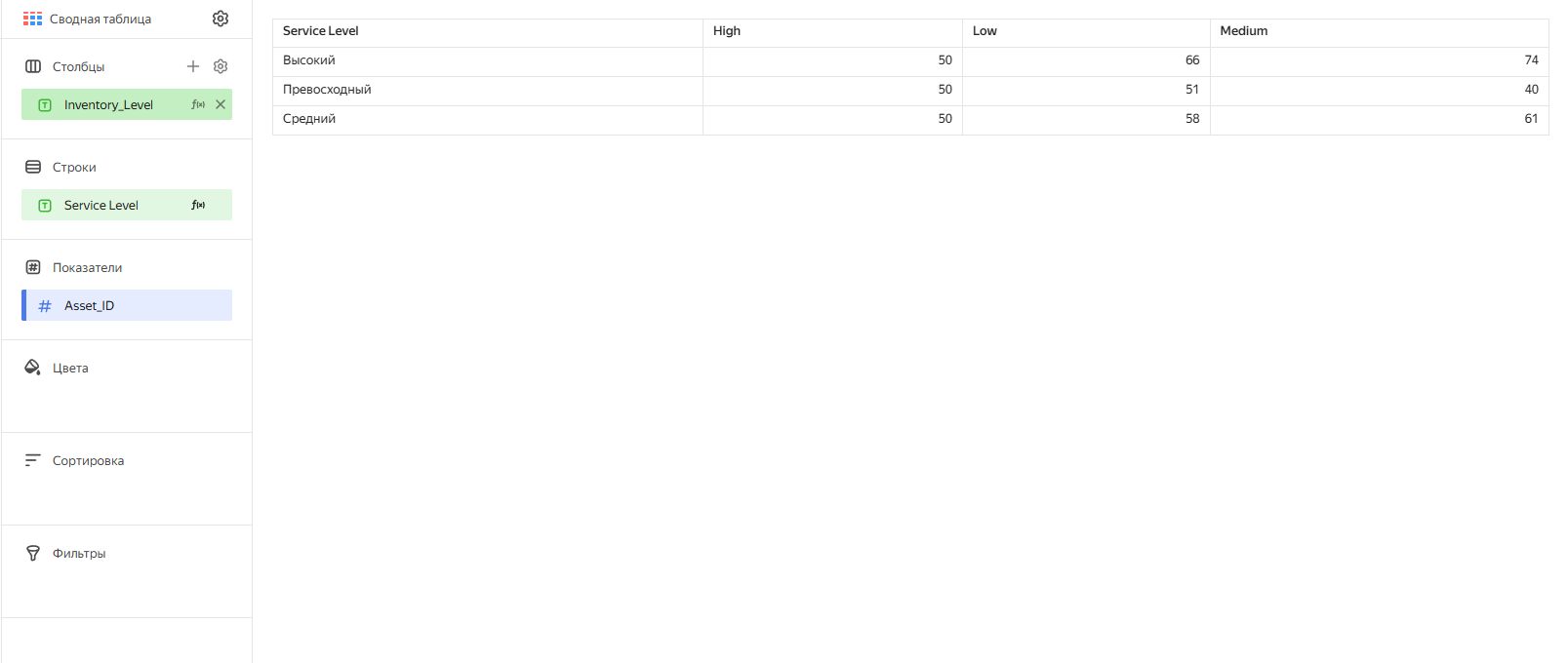
**Анализ запасов и сервиса (Сводная таблица):**

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

**Настройка:**

Строки: Уровень обслуживания (Service Level)

Столбцы: Inventory\_Level

Показатель: COUNT([Asset\_ID])

**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%203.1](https://github.com/youngvoyt/DEP-MGPU/tree/ccc21610617262d270c1bcb984f4a608465f5077/Module03/Lab%203.1)]