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

**Лабораторная работа № 3.1**

**Вариант 30**

по дисциплине «Платформы Data Engineering»

Выполнил:

студент группы БД-251м

Направление подготовки/Специальность

38.04.05 - Бизнес-информатика

Трухачев Никита Алексеевич

(Ф.И.О.)

Проверил:

Доцент, к.т.н

(ученая степень, звание)

Босенко Тимур Муртазович

(Ф.И.О.)

Москва 2025

**Лабораторная работа 3.1 Создание интерактивного аналитического дашборда наоснове витрин данных**

**1. Введение**

**Постановка цели**

Цель проекта — создание системы аналитики для оценки и оптимизации **затрат, связанных с системами хранения и обработки данных** внутри компании.  
На основе данных внутреннего опроса сотрудников была построена база данных и разработаны **аналитические витрины**, используемые для построения интерактивных дашбордов в **Yandex DataLens**.

**Выбранная тема**

Тема проекта — *управление затратами в системах работы с данными*, включая анализ прозрачности процессов, автоматизации мониторинга, распределения бюджета и зрелости управления.

**Целевая аудитория опроса**

* Специалисты и руководители IT, аналитических, финансовых и продуктовых подразделений.
* Цель опроса — собрать оценку текущего состояния управления затратами, определить узкие места и зоны для оптимизации.
* Ссылка на [**опрос**](https://github.com/ForestGnom/DEP-MGPU/blob/main/%D0%9B%D0%91_3_1/docs/survey/%D0%9E%D0%BF%D1%80%D0%BE%D1%81.pdf)
* **2. Анализ результатов**

На основе данных опроса были созданы визуализации в Yandex DataLens, отражающие взаимосвязь между типами систем, уровнем зрелости, эффективностью и отклонениями бюджета.

**Ключевые визуализации (всего 3 графика)**

**1. Линейный график "Типы систем, используемые в подразделении"**

**Источник данных:** analytics.mart\_company\_strategy\_metrics

* **Ось X:** department (департаменты), systems\_used (типы систем)
* **Ось Y:** total\_records (количество ответов)
* **Цвета:** systems\_used (типы систем)
* **Интерпретация:** показывает распределение стратегий и систем по департаментам, помогает оценить эффективность различных технологий и практик.

**2. Линейный график "Уровень зрелости управления затратами"**

**Источник данных:** analytics.mart\_maturity\_summary

* **Ось X:** maturity\_level (уровень зрелости)
* **Ось Y:** total\_records (количество ответов)
* **Интерпретация:** демонстрирует соотношение зрелости управления затратами и числа подразделений на каждом уровне.

**3. Круговая диаграмма "Средняя доля затрат на хранение данных внутри департаментов"**

**Источник данных:** analytics.mart\_department\_summary

* **Метрика:** avg\_storage\_share (средняя доля затрат на хранение данных)
* **Группировка:** department (департаменты)
* **Интерпретация:** отображает относительное распределение затрат на хранение данных между департаментами компании.
* **3. Формирование витрин**

**Источник данных**

Все витрины формируются на основе таблицы raw\_data.cost\_management\_survey с использованием агрегатных функций и аналитических выражений.

**Логика расчета ключевых метрик**

**Эффективность управления затратами:**

CASE

WHEN AVG(budget\_deviation\_pct) < -5 THEN 'Высокая эффективность'

WHEN AVG(budget\_deviation\_pct) BETWEEN -5 AND 5 THEN 'Средняя эффективность'

ELSE 'Низкая эффективность'

END AS cost\_management\_effectiveness,

**Категоризация систем**

CASE

WHEN AVG(budget\_deviation\_pct) < -5 THEN 'Высокая эффективность'

WHEN AVG(budget\_deviation\_pct) BETWEEN -5 AND 5 THEN 'Средняя эффективность'

ELSE 'Низкая эффективность'

END AS cost\_management\_effectiveness,

**Рекомендации по стратегии**

CASE

WHEN AVG(budget\_deviation\_pct) > 10 THEN 'Требуется пересмотр стратегии управления затратами'

WHEN AVG(budget\_deviation\_pct) BETWEEN 0 AND 10 THEN 'Оптимизация возможна'

WHEN AVG(budget\_deviation\_pct) < 0 THEN 'Эффективные практики — можно масштабировать'

ELSE 'Поддерживать текущий уровень'

END AS strategic\_recommendation

* **4. DataLens Дашборд**

**Дашборд "Управление затратами в системах работы с данными"**

**Ссылка:** [Yandex DataLens Dashboard](https://datalens.yandex/v6rumrm47eg4g)

Включает три ключевых визуализации:

1. **Линейный график "Типы систем, используемые в подразделении"** — распределение стратегий и систем по подразделениям.
2. **Линейный график "Уровень зрелости управления затратами"** — динамика зрелости управления затратами.
3. **Круговая диаграмма "Средняя доля затрат на хранение данных внутри департаментов"** — доли затрат на хранение данных по подразделениям.

* **5. Выводы**

1. **Затраты на хранение данных варьируются между департаментами**, при этом Операции/Бизнес-подразделение демонстрируют наибольшую долю затрат (32,16%), что указывает на потенциал оптимизации.
2. **Уровень зрелости управления затратами значительно различается по компании**, с диапазоном от начального до установленных процессов, что требует стандартизации подходов.
3. **Аналитические подразделения наиболее активно используют различные системы**, что коррелирует с их ролью в работе с данными, но требует контроля эффективности затрат.
4. **Наблюдается значительный разброс в эффективности управления затратами**, с показателями от 150 до 1920 единиц, что свидетельствует о необходимости улучшения процессов.
5. **Основные зоны роста** — оптимизация распределения затрат на хранение между департаментами и повышение стабильности процессов управления.

* **6. Заключение**

Создана аналитическая система, объединяющая опросные данные, SQL-витрины и визуализацию в Yandex DataLens.  
Результаты позволяют:

* Повысить прозрачность распределения затрат;
* Сопоставить эффективность различных систем и практик;
* Поддержать принятие решений на основе данных.