**PET-проект: HR аналітика**

Документація та презентація проекту "HR аналітика" будуть ключовими інструментами для висвітлення результатів аналізу та оцінки ефективності управління персоналом. У документації буде детально описано методологію аналізу, використані показники та моделі, а також виявлені основні висновки та рекомендації. Презентація включатиме в себе візуальні матеріали, такі як графіки та діаграми, що ілюструють знайдені закономірності та тренди.

**Мета проекту:** створення аналітичного звіту для оцінки ефективності управління персоналом.

**Обгрунтування актуальності теми:** дана тема має актуальність у бізнес середовищі, так як може допомогти компаніям в оптимізації ресурсів, підвищення ефективності процесів, підвищення конкурентоспроможності та підвищення рівня задоволеності працівників.

**Методологія проекту**

**Тема проекту** обрана на основі актуальності для бізнесу, так як основний ресурс кожної компанії – люди, які працюють у ній. Дана тема допоможе розкрити більш глибоке розуміння HR ресурсів у компанії та ефективність роботи з ними.

**Використані технології**: у даній роботі використано EXCEL, BigQuery та SQL, Tableau Public.

EXCEL було використано для підготовки даних до аналізу та первинного аналізу.

BigQuery як хмарний сервіс Google було використано з метою здійснення необхідних SQL запитів.

Tableau Public було використано з метою візуалізації результатів.

**Опис проекту:**

Проект спрямований на аналіз різних аспектів управління персоналом в організації. Основні напрямки проекту включають:

1. Статистика по персоналу:

* Розподіл працівників за статтю та віком для отримання детального портрета персоналу.
* Аналіз розподілу працівників за департаментами, відділами та позиціями для оцінки внутрішньої структури та тенденцій.

1. Статистика звільнень та плинності персоналу:

* Визначення кількості звільнених працівників за період для оцінки динаміки звільнень персоналу.
* Аналіз причин звільнень за termination type, зосереджений на department type.
* Розрахунок turnover rate для оцінки стабільності кадрового складу.

1. Аналіз ефективності найму:

* Оцінка середньої тривалості перебування працівника у компанії на основі дат початку та завершення роботи.

1. Аналіз ефективності роботи та задоволення працівників:

* Вивчення оцінок працівників за різними показниками, такими як Satisfaction Score та Work-Life Balance Score.
* Аналіз різниці в ефективності та задоволенні працівників за різними демографічними групами для ідентифікації потенційних тенденцій та виявлення факторів, що впливають на результативність та задоволеність.

1. Розподіл працівників за типами та класифікацією:

* Аналіз розподілу працівників за типом зайнятості (Contract, Full-time, Part-time) для виявлення тенденцій.

При пошуці датасету для власного проекту я здійснила аналіз різноманітних джерел з яких я можу отримати дану інформацію. Мені необхідно було знайти HR or Recruitment data, які міститимуть достатньо інформації для аналізу даних. Я розглядала можливість використання штучного інтелекту для створення даних, однак даний підхід виявився невдалим, так як після детальної розробки даних, що повинні міститись у таблиці виходячи з recruitment metrics, я отримала дані з якими не можна було працювати, так як не були релеватними саме дати у створеній таблиці.

Серед усіх джерел я обрала саме Kaggle, так як дана платформа містила найбільш цікаві для мене датасети. Я обрала наступний датасет: <https://www.kaggle.com/datasets/ravindrasinghrana/employeedataset?resource=download>

У даному датасеті міститься 4 таблиці, однак я обрала лише дві, які скомпонувала між собою. Цей вибір зумовлений тим, що всі таблиці між собою не поєднувались, так як містили різні дані.

Таблиця Employee\_Data містить наступні колонки:

1. **Employee ID:** Unique identifier for each employee in the organization.
2. **First Name:** The first name of the employee.
3. **Last Name:** The last name of the employee.
4. **Start Date:** The date when the employee started working for the organization.
5. **Exit Date:** The date when the employee left or exited the organization (if applicable).
6. **Title:** The job title or position of the employee within the organization.
7. **Supervisor:** The name of the employee's immediate supervisor or manager.
8. **Email:** The email address associated with the employee's communication within the organization.
9. **Employee Status:** The current employment status of the employee (e.g., Active, On Leave, Terminated).
10. **Employee Type:** The type of employment the employee has (e.g., Full-time, Part-time, Contract).
11. **Employee Classification Type:** The classification type of the employee (e.g., Exempt, Non-exempt).
12. **Termination Type:** The type of termination if the employee has left the organization (e.g., Resignation, Layoff, Retirement).
13. **Termination Description:** Additional details or reasons for the employee's termination (if applicable).
14. **Department Type:** The broader category or type of department the employee's work is associated with.
15. **Division:** The division or branch of the organization where the employee works.
16. **DOB (Date of Birth):** The date of birth of the employee.
17. **State:** The state or region where the employee is located.
18. **Job Function:** A brief description of the employee's primary job function or role.
19. **Gender Code:** A code representing the gender of the employee (e.g., M for Male, F for Female, N for Non-binary).
20. **Performance Score:** A score indicating the employee's performance level (e.g., Excellent, Satisfactory, Needs Improvement).
21. **Current Employee Rating:** The current rating or evaluation of the employee's overall performance.
22. **Engagement Score:** A calculated numerical score representing the level of employee engagement based on survey responses.
23. **Satisfaction Score:** A numerical score indicating employee satisfaction with various aspects of their job and workplace.
24. **Work-Life Balance Score:** A numerical score reflecting employee perceptions of the balance between work and personal life.

**План проекту**

Етапи та задачі

Перший етап: Розробка даних та підготовка даних до аналізу

* Пошук даних.
* Форматування даних.
* Очищення даних.

Другий етап: Попередній аналіз даних.

* Вивчення закономірностей та тенденцій.
* Визначення даних, які необхідно отримати за допомогою написання SQL запитів.
* Завантаження табличних даних у BigQuery.
* Написання SQL запитів.

Третій етап: Візуалізація даних.

* Завантаження даних до Tableau Public.
* Робота над візуалізаціями.
* Розробка дашбордів.

Четвертий етап: Формування висновків та представлення даних.

* Формування висновків на основі отриманих даних та візуалізацій.
* Створення презентації.

У ході першого етапу після отримання даних я здійснила наступні кроки:

* Спочатку дані було розділено по стовпчиках за роздільником
* Форматувала дані у табличні
* За допомогою PowerQuery здійснила очищення від недійсних значень
* Здійснила також форматування даних, зокрема, було також змінено дати у правильний формат дат та ін.

На першому етапі я вивчила дані і спробувала проаналізувати можливі закономірності, що можуть виникнути. Також написала опис, що саме хотіла б проаналізувати у ході даного проекту. Даний опис міститься у «Опис проекту» на початку.

Після такого аналізу мені необхідно було отримати наступні дані:

1. Вік кожного працівника у компанії.
2. Кількість місяців/років досвіду у кожного працівника по департаменту, відділу враховуючи те, що деякі працівники вже не працюють у компанії.

Після цього я завантажила таблицю Employee\_Data до BigQuery та написала наступні запити для отримання необхідних даних:

SELECT EmpoyeeID, DATE\_DIFF(DATE('2023-09-01'), DOB, YEAR) AS Age

FROM `able-involution-403316.EMPLOYEE\_DATA.EMPLOYEE\_DATA`;

SELECT EmpoyeeID, Division, DepartmentType,

    DATE\_DIFF(IFNULL(ExitDate, CURRENT\_DATE()), StartDate, MONTH) AS month\_worked

FROM `able-involution-403316.EMPLOYEE\_DATA.EMPLOYEE\_DATA`

WHERE ExitDate IS NOT NULL;

SELECT EmpoyeeID, Division, DepartmentType,

   DATE\_DIFF(IFNULL(ExitDate, CURRENT\_DATE()), StartDate, YEAR) AS years\_worked

FROM `able-involution-403316.EMPLOYEE\_DATA.EMPLOYEE\_DATA`;

**Аналіз даних**

**Розподіл працівників за статтю та віком**: використання дескриптивної статистики для розрахунку середнього віку працівників та розподіл за статтю для отримання детального портрета персоналу.

**Розподіл працівників за департаментами, відділами та позиціями**: використання категоріального аналізу для оцінки кількості працівників за внутрішньою структурою організації.

**Кількість звільнених працівників за період**: використання показника частоти для оцінки динаміки звільнень персоналу та ідентифікації тенденцій.

**Аналіз причин звільнень за termination type**: використання крос-табуляції та статистичних тестів для вивчення зв'язку між причинами звільнень та внутрішньою структурою організації.

**Turnover rate**: використання розрахунку показника Turnover rate для оцінки стабільності кадрового складу та виявлення тенденцій у зміні персоналу.

**Візуалізація даних**

Наступним етапом стало завантаження даних до Tableau Public та створення графіків та дашбордів.

Всього було створено 3 дашборди та 14 графіків.

Перший дашборд демонструє Quantity of employees in the company by parameters.

У даний дашборд входить 4 візуалізації.

Перший графік - bar - кількість працівників на кожній позиції за гендерним розподілом.

Другий графік - bar - середній вік працівників по кожній позиції.

Третій графік - bar - кількість працівників у кожному департаменті за гендерним розподілом.

Четвертий графік - bar - кількість працівників у кожному відділі за гендерним розподілом.

Другий дашборд демонструє Terminations analysis. У даному дашборді містяться 3 візуалізації.

Перша візуалізація – pie chart – Terminations by type. На ньому ми бачимо кількість працівників, які припинили роботу в компанії по одній з 4 причин: Voluntary, Involuntary, Resignation, Retirement.

Другий графік – bar - кількість звільнень у кожному департаменті.

Третій графік – line - демонструє тенденцію до припинення роботи в компанії.

Третій дашборд демонструє Satisfaction/work-life balance scores by parameters.

У даний дашборд входить 4 візуалізації.

Перший графік – line - Satisfaction/work-life balance by age – показує середню оцінку рівня задоволеності роботою та рівня балансу роботи та особистого життя залежно від віку.

Друга візуалізація – pie chart – Satisfaction by departments. Показує середню оцінку рівня задоволеності роботою по кожному департаменту.

Третя візуалізація – pie chart – Work-life balance by departments. Показує середню оцінку балансу між роботою та особистим життям по кожному департаменту.

Четверта візуалізація – map – показує мапу США і кількість працівників у кожному штаті з розподілом за департаментами.

Візуалізація 1 - bar - Average employee tenure – показує середню кількість років перебування у компанії працівників за департаментами.

Візуалізація 2 – bar - Quantity of current employees by age – відповідно до даного дашборду ми можемо відслідкувати кількість працівників, які наразі у компанії, залежно від віку.

Візуалізація 3 – bar - Type of employment contracts – даний дашборд демонструє кількість працівників залежно від типу контракту та класифікації типу працевлаштування.

Для обчислення Turnover rate потрібно поділити кількість звільнених працівників на кількість всіх працівників за певний період. У нашому випадку це весь період роботи компанії. Та помножити на 100.

Turnover rate – 51.1 %

**Висновки та рекомендації для бізнесу**

Гендерне співвідношення працівників, які наразі продовжують працювати, у кожному департаменті, відділу та позиції не є завжди рівним. Зокрема, відділи Production 392 – чоловіків, 614 – жінок; Sales 105 - чоловіків, 62 - жінки.

* Рекомендація: звернути увагу на даний розподіл і причини яким він зумовлений для забезпечення гендерної рівності.

Аналіз тенденції до звільнень у компанії враховуючи розподіл по типу звільнень. У компанії існує 4 причини припинення роботи: Voluntary – 388 працівників, Involuntary – 388 працівників, Resignation – 380 працівників, Retirement – 377 працівників за весь час функціонування компанії. Причини припинення роботи розподілені рівномірно.

Найбільше звільнень саме у департаменті Production – 2020, а найменше у Executive Office - 24. Однак це корелює з розмірами департаментів. На час останніх записів (серпень 2023 рік) у департаменті Production - 1006 працівників, а у Executive Office – 5 працівників.

Тенденції звільнень показані з листопаду 2018 року і по серпень 2023 року включно. Спостерігаємо тенденцію до збільшення кількості звільнень протягом всього періоду. У липні 2023 року найвища кількість працівників, які припинили роботу у компанії - 127. Наступний місяць – серпень, демонструє різке відхилення – 26 звільнень.

При аналізі найбільшої кількості звільнень у липні, найчастіше причиною припинення роботи у компанії є Voluntary - 47, а департамент з якого звільнялись найчастіше - Production – 93 працівника (з них 35 – Voluntary).

* Рекомендація: додатково збирати причини звільнень з більшою деталізацією. Рекомендуємо запровадити exit interview на яких отримувати відгуки працівника щодо роботи компанії, причин звільнень та отримати оцінку окремих показників для аналізу бренду роботодавця. Звернути увагу на збільшення кількості звільнень кожного періоду.

Середня оцінка Satisfaction та work-life balance score – 3. За розподілом по департаментах оцінка є рівномірною. За розподілом за віком середня оцінка work-life balance є рівномірною, а Satisfaction має незначні відмінності. Найвищий рівень задоволеності роботою у працівників 40 років, а найнижчий у працівників 80 років.

У департаменті Sales найнижчий рівень задоволеності роботою у наступних штатах– NV (1), MT (2), OH (2), CT (2), NH (2). Production – RI (2).

* Рекомендація: звернути увагу на штати у яких критично низький Satisfaction score. Середня оцінка по компанії є доволі низькою і тому радимо здійснити додаткове опитування з детальними поясненнями для підвищення рівня задоволеності працівників.

3,5 років – середня кількість перебування працівника у компанії. Даний розподіл здійснено по департаментах і значних відхилень залежно від департаменту не спостерігається.

Середній вік працівників по відділах від 43 років до 57 років.

Відповідно до графіку розподілу віку працівників, найбільша кількість працівників, які наразі працюють у компанії, віком 30 років (138 працівників), 70 років (138 працівників) та 61 рік (131 працівник).

У компанії рівномірний розподіл працівників, які працюють контрактом, повний робочий день та неповний робочий день.

Turnover rate – 51.1 %