

Đề bài chủ đề HR

Employee Turnover & Attrition Analysis – Phân tích Biến động và Tỷ lệ Nghỉ việc Nhân viên

Thử thách dữ liệu thực chiến thuộc cuộc thi **NextGen Analytics Challenge 2025**, dành cho bất kỳ ai yêu thích và đam mê lĩnh vực dữ liệu – đặc biệt là ứng dụng trong ngành Nhân sự.

A. Nội dung đề bài

Chủ đề HR

Employee Turnover & Attrition Analysis – Phân tích Biến động và Tỷ lệ Nghỉ việc Nhân viên

Bối cảnh bài toán, vai trò và mục tiêu phân tích

Bối cảnh

Bạn đang là **Chuyên viên Phân tích Dữ liệu Nhân sự (HR Data Analyst)** thuộc **Phòng Chiến lược & Quản trị Nguồn nhân lực** của một tập đoàn đa ngành quy mô lớn.

Ban lãnh đạo công ty đang đối mặt với tỷ lệ nghỉ việc (Attrition) gia tăng trong 2 năm qua, ảnh hưởng đến:

- Chi phí tuyển dụng, đào tạo lại đội ngũ
- Tinh thần nhân viên còn lại
- Khả năng duy trì hiệu suất kinh doanh

Công ty đã thu thập một bộ dữ liệu chi tiết về đặc điểm hồ sơ nhân viên, điều kiện làm việc, hiệu suất và mức độ gắn bó. Ban lãnh đạo yêu cầu bạn **phân tích và trực quan hóa dữ liệu** nhằm làm rõ các nguyên nhân dẫn đến nghỉ việc và đề xuất giải pháp quản trị rủi ro nhân sự.

Vai trò của bạn

Bạn đóng vai trò:

Chuyên viên phân tích dữ liệu nhân sự (HR Data Analyst).

Bạn chịu trách nhiệm:

- Làm sạch dữ liệu, trực quan hóa và phân tích các biến liên quan đến khả năng nghỉ việc của nhân viên.
- Xây dựng **dashboard Power BI** trực quan, dễ sử dụng cho ban lãnh đạo và bộ phận nhân sự ra quyết định.
- Tóm tắt các phát hiện quan trọng và đưa ra **đề xuất cải thiện tỷ lệ gắn bó** dựa trên kết quả phân tích.

Người xem báo cáo:

- Ban Lãnh đạo Cấp cao (CEO, CHRO - Giám đốc Nhân sự, CFO)
- Phòng Nhân sự Chiến lược và Tuyển dụng
- Quản lý các khối phòng ban

Mục đích báo cáo

Báo cáo và dashboard Power BI cần trả lời **các câu hỏi cụ thể và nhất quán**, phục vụ mục tiêu:

- **Hiểu rõ đặc điểm hồ sơ nhân viên:**
 - Nhân khẩu học (tuổi, giới tính, trình độ học vấn, tình trạng hôn nhân).
 - Thông tin công việc (phòng ban, cấp bậc, thu nhập, thâm niên).
 - Điều kiện làm việc (Work-Life Balance, Overtime, Recognition...).
- **Đo lường và trực quan hóa tỷ lệ nghỉ việc:**
 - Tỷ lệ nghỉ việc (Attrition) **chung toàn công ty**.
 - So sánh tỷ lệ nghỉ việc giữa **các nhóm nhân viên khác nhau**.
- **Xác định yếu tố rủi ro chính:**
 - Biến nào liên quan mạnh nhất đến khả năng nghỉ việc?
 - Những nhóm nhân viên nào có rủi ro cao hơn mức trung bình?
- **Hỗ trợ ra quyết định:**
 - Đề xuất **nhóm nhân viên ưu tiên giữ chân và cải thiện điều kiện làm việc**.

- Đề xuất **chính sách hỗ trợ, ghi nhận, điều chỉnh công việc** để giảm tỷ lệ nghỉ việc.

Các yếu tố cần đạt được sau khi phân tích

Để đảm bảo tất cả thí sinh dự thi cùng trả lời một vấn đề thống nhất, báo cáo bắt buộc phải bao gồm:

- **Tổng quan toàn công ty:**
 - Số lượng nhân viên
 - Tỷ lệ nghỉ việc trung bình (%)
 - Phân bố theo phòng ban, cấp bậc công việc
- **So sánh đặc điểm nhóm "Stayed" vs "Left":**
 - Biểu đồ hoặc bảng thể hiện sự khác biệt về:
 - Tuổi, giới tính, trình độ học vấn
 - Thu nhập trung bình
 - Số lần được thăng chức
 - Khoảng cách từ nhà đến công ty
- **Phân tích yếu tố ảnh hưởng đến nghỉ việc:**
 - Tương quan tỷ lệ nghỉ việc với các yếu tố:
 - Work-Life Balance
 - Job Satisfaction
 - Overtime
 - Performance Rating
 - Company Reputation & Recognition
- **Xác định 3 yếu tố rủi ro lớn nhất:**
 - Nêu rõ lý do lựa chọn, kèm visual minh họa
- **Đề xuất giải pháp:**
 - Tối thiểu 3 kiến nghị nghiệp vụ (ví dụ: cải thiện Work-Life Balance, chính sách ghi nhận, hỗ trợ làm thêm giờ...)
- **Dashboard Power BI:**
 - Trang tổng quan: KPI chính và phân bố nhân viên
 - Trang phân tích chi tiết: bộ lọc (slicer) cho phòng ban, cấp bậc, trạng thái làm việc

- Tương tác (drill-through, highlight)
- Giao diện rõ ràng, trực quan, dễ đọc

Mô tả Dataset

Bộ dữ liệu này được xây dựng mô phỏng, phục vụ mục đích **phân tích và dự báo tình trạng nghỉ việc (attrition)** của nhân viên. Mỗi bản ghi đại diện cho một hồ sơ nhân viên, bao gồm thông tin chi tiết về:

- Nhân khẩu học (tuổi, giới tính, trình độ học vấn)
- Đặc điểm công việc (phòng ban, vị trí, mức lương, thâm niên)
- Các yếu tố liên quan đến sự gắn kết và mức độ hài lòng

Bộ dữ liệu chứa tổng cộng **59.598 dòng dữ liệu** và được chia sẵn thành các phần dùng cho huấn luyện và kiểm tra mô hình. Tuy nhiên, trong cuộc thi này, thí sinh sẽ tập trung vào **khai thác, trực quan hóa và phân tích dữ liệu thực trạng** thay vì xây dựng mô hình dự báo.

Mục tiêu chính của việc phân tích là giúp người xem **hiểu rõ các yếu tố ảnh hưởng đến nguy cơ nghỉ việc**, nhận diện nhóm nhân viên rủi ro cao và xây dựng nền tảng dữ liệu hỗ trợ ra quyết định chiến lược về quản trị nhân sự và duy trì đội ngũ.

Thí sinh sẽ sử dụng Power BI để trực quan hóa dữ liệu, trình bày các xu hướng, khám phá các mối liên hệ quan trọng và tạo dashboard tương tác phục vụ ban lãnh đạo và bộ phận nhân sự.

Bảng chú giải dữ liệu

#	Tên biến	Kiểu dữ liệu	Mô tả
1	Employee ID	Số nguyên	Mã định danh duy nhất của nhân viên
2	Age	Số nguyên	Tuổi của nhân viên
3	Gender	Chuỗi	Giới tính nhân viên
4	Years at Company	Số nguyên	Số năm nhân viên làm việc tại công ty

#	Tên biến	Kiểu dữ liệu	Mô tả
5	Job Role	Chuỗi	Phòng ban/ vị trí công việc
6	Monthly Income	Số nguyên	Thu nhập hàng tháng của nhân viên
7	Work-Life Balance	Chuỗi	Đánh giá cân bằng công việc và cuộc sống
8	Job Satisfaction	Chuỗi	Mức độ hài lòng công việc
9	Performance Rating	Chuỗi	Đánh giá hiệu suất nhân viên
10	Number of Promotions	Số nguyên	Số lần nhân viên được thăng chức
11	Overtime	Chuỗi	Nhân viên có làm thêm giờ không
12	Distance from Home	Số nguyên	Khoảng cách từ nhà nhân viên đến công ty
13	Education Level	Chuỗi	Trình độ học vấn cao nhất
14	Marital Status	Chuỗi	Tình trạng hôn nhân
15	Number of Dependents	Số nguyên	Số người phụ thuộc
16	Job Level	Chuỗi	Cấp bậc công việc
17	Company Size	Chuỗi	Quy mô công ty mà nhân viên làm việc
18	Company Tenure	Số nguyên	Số năm kinh nghiệm của nhân viên trong ngành
19	Remote Work	Chuỗi	Nhân viên có làm việc từ xa hay không

#	Tên biến	Kiểu dữ liệu	Mô tả
20	Leadership Opportunities	Chuỗi	Nhân viên có cơ hội lãnh đạo không
21	Innovation Opportunities	Chuỗi	Nhân viên có cơ hội đổi mới không
22	Company Reputation	Chuỗi	Cảm nhận của nhân viên về danh tiếng công ty
23	Employee Recognition	Chuỗi	Mức độ ghi nhận nhân viên
24	Attrition	Chuỗi	Biến mục tiêu - nhân viên đã nghỉ việc hay chưa

Lưu ý: Một số cột trong dataset có thể chứa giá trị bất thường, thí sinh được khuyến khích kiểm tra và xử lý dữ liệu trước khi xây dựng báo cáo phân tích.

B. Hướng dẫn kết nối dữ liệu

Cài đặt Power BI

Thí sinh cần cài đặt Power BI Desktop để làm bài thi.

→ Truy cập link sau để xem hướng dẫn chi tiết cài đặt:

https://datapot.vn/student-support-portal/#Doi_voi_Power_BI

Hướng dẫn kết nối dữ liệu từ SQL Server của Datapot

→ Truy cập link sau để xem hướng dẫn kết nối vào SQL Server:

https://datapot.vn/power-bi-moc-lab-1-prepare-data-in-power-bi-desktop/#Task_2_Lay_du_lieu_tu_SQL_Server

Thông tin kết nối server

Connection type: Microsoft SQL Server

Server: analytics-challenge-sql.datapot.edu.vn

Authentication: SQL Server Database

User name: nextgen_user

Password: Nextgen2025_01

C. Thông tin cuộc thi

Datapot phối hợp cùng các đối tác chuyên môn tổ chức cuộc thi **NextGen Analytics Challenge 2025**, với mục tiêu tạo ra một sân chơi thực tế để người học có cơ hội rèn luyện kỹ năng xử lý – phân tích – trực quan hóa dữ liệu thông qua các bài toán mô phỏng từ doanh nghiệp.

Cuộc thi kéo dài từ **tháng 7 đến tháng 12 năm 2025**, bao gồm **4 chặng thi chính**, tương ứng với **4 lĩnh vực phân tích**:

- Tài chính (Finance)
- Nhân sự (HR)
- Chuỗi cung ứng (Supply Chain) & Hậu cần (Logistics)
- Bán lẻ (Retail) & Người tiêu dùng (Consumer)

Thông tin chung về cuộc thi

- **Form nộp bài:** <https://forms.gle/c7JjRWhtNgbtYSyT6>
- **Website Datapot Challenge | Hall of Fame:** <https://datapot.vn/challenge-hall-of-fame/>

Kênh liên lạc – cập nhật thông tin chính thức

- **Fanpage Datapot:** <https://www.facebook.com/DatapotAnalytics>
- **Group Facebook DABI – Data Analytics and Business Intelligence Vietnam:** <https://www.facebook.com/groups/nguoiламdata>

Kênh hỗ trợ chuyên môn

- **Zalo Group hỗ trợ Q&A:** <https://zalo.me/g/ekxslq015>

- **Kênh YouTube Datapot:** <http://info.datapot.vn/kenhYouTubeDatapot>
- **Tài nguyên hỗ trợ thí sinh:** <https://datapot.vn/tong-hop-tai-nguyen-ho-tro-thi-sinh-nextgen-analytics-challenge-2025/>