DMI Advanced HR Analytics

Magdalena Pietrzak, Hubert Domagała Michał Głąb

1. **Cele i założenia projektu:**

WAIT posiada informacje o deklarowanych kompetencjach osób gotowych do pracy w projektach, ale brakuje wiedzy na temat ewolucji tych kompetencji. Wyzwanie polega na analizie zebranych danych w celu wyciągnięcia użytecznych wniosków, które pomogą w tworzeniu efektywnych zespołów projektowych. Dodatkowo, konieczne jest przeprowadzenie segmentacji członków WAIT poprzez grupowanie i klastrowanie.

Cele projektu:

**Analiza danych kompetencyjnych:** Wyciągnięcie użytecznych wniosków z dotychczas zebranych danych dotyczących kompetencji członków WAIT.

**Segmentacja członków:** Przeprowadzenie segmentacji i klastrowania członków WAIT, aby zidentyfikować "rozróżnialne" grupy użytkowników.

**Zrozumienie społeczności:** Na podstawie dostępnych danych sformułowanie interesujących wniosków dotyczących społeczności WAIT, które mogą wspierać konstrukcję zespołów projektowych.

Te cele pozwolą lepiej zrozumieć strukturę i dynamikę kompetencji w społeczności WAIT, co w efekcie przyczyni się do bardziej efektywnego zarządzania zasobami ludzkimi w projektach.

1. **Docelowy odbiorca projektu:**

Osoby zarządzające projektami w WAIT lub ludzie poszukujący innych ludzi do swoich projektów

1. **Zasoby i ograniczenia:**

**Ograniczenia kompetencyjne**

Większość osób realizujących projekt nie ma doświadczenia w tego typu przedsięwzięciach i dopiero zaczyna naukę podstaw uczenia maszynowego. Brak wcześniejszego doświadczenia może wpływać na tempo pracy i jakość realizacji zadań.

**Ograniczenie czasowe**

Projekt musi zostać ukończony do 23 czerwca 2024 roku, co stanowi ostateczny termin zakończenia prac. Ograniczony czas wymaga efektywnego zarządzania zasobami i ścisłego przestrzegania harmonogramu.

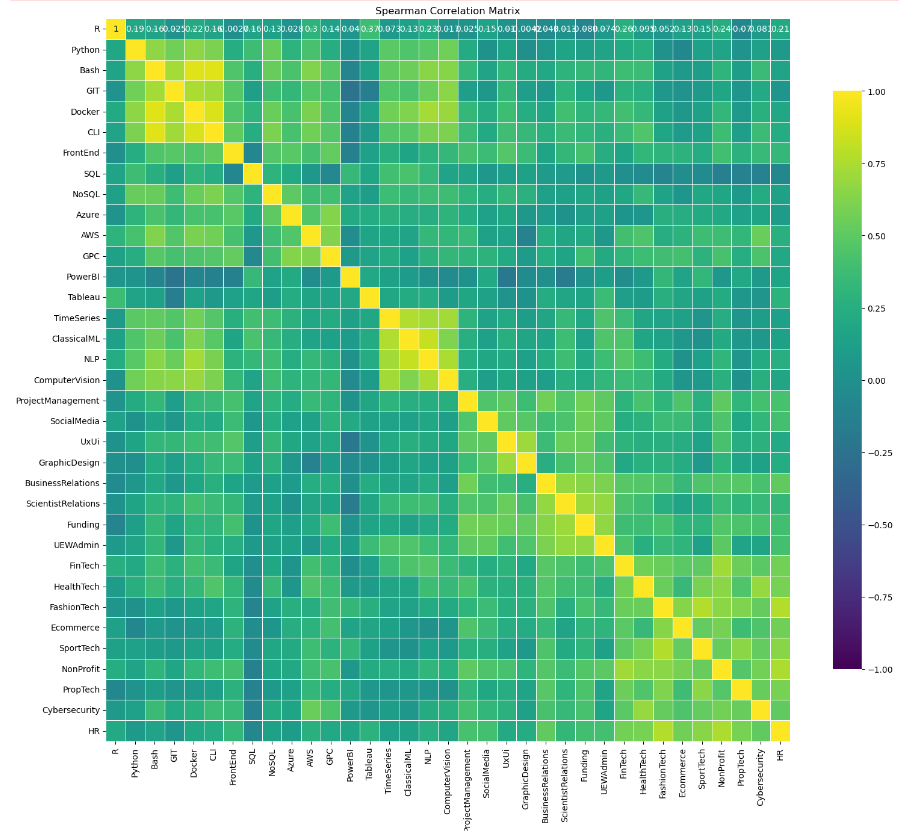
1. **Analiza wyników**

Poniższy obrazek przedstawia macierz korelacji Spearmana, która jest graficzną reprezentacją współczynników korelacji rang Spearmana między różnymi zmiennymi. Macierz jest kwadratowa, a obie osie są oznaczone różnymi kompetencjami, takimi jak Python, R, SQL, Excel i inne.

Kolory na mapie ciepła reprezentują różne wartości współczynnika korelacji. Ciemnoniebieski oznacza współczynnik korelacji bliski -1, co wskazuje na silną negatywną korelację między dwoma kompetencjami. To oznacza, że gdy jedna kompetencja jest wysoko oceniana, druga jest zwykle nisko oceniana, i na odwrót.

Zielony kolor reprezentuje małą lub brak korelacji, co oznacza, że rankingi dwóch kompetencji nie są ze sobą powiązane.

Ciemnożółty kolor oznacza współczynnik korelacji bliski 1, co wskazuje na silną pozytywną korelację. To oznacza, że dwie kompetencje są zwykle oceniane podobnie - gdy jedna jest wysoko oceniana, druga również jest zwykle wysoko oceniana.

****

Analizując macierz, można zauważyć pewne interesujące zależności. Na przykład, kompetencje Python i R wydają się być silnie skorelowane, co ma sens, ponieważ obie są popularnymi językami programowania używanymi w analizie danych. Z drugiej strony, Python i Excel wydają się być słabo skorelowane, co może wynikać z faktu, że Excel jest częściej używany do prostszych zadań analizy danych, podczas gdy Python jest preferowany do bardziej złożonych analiz.

Największe korelacje widać jednak w czterech skupionych grupach kompetencji:

* Fintech, Heatech, Fashiontech, Ecomemerance, Sportech, NonProfit, Cyberseciurity oraz Hr;
* UEWadmin, Funding, ScientisRelations, BuiesnessRelations, GrephicDesign;
* ProjectMenager, ComputerVision, NLP, ClassicalML
* Python, Bash, Git, Docker, CLI