Data Wrangling Übung 1

Die Beste Gemeinde des Kantons

1) Idee

Als unabhängiges Forschungsinstitut haben Sie von einem Kanton den Auftrag erhalten, eine Rangliste der einzelnen Gemeinden zu erstellen. Einer der Gemeinderäte hat aber Wind davon bekommen und «ermutigt» Sie, die Daten so hinzubiegen, dass seine Gemeinde möglichst gut dasteht und seine Nachbarsgemeinden möglichst schlecht.

2) Vorbereitung

Tragen Sie sich bis am **2. März** in folgender Umfrage ein und suchen Sie sich einen Kanton aus. https://docs.google.com/spreadsheets/d/10G5LzjpyxfxwwCCBCkYsnrq-5rSYgKusuuXvQBWemCk/edit?usp=sharing (Pro Kanton maximal 1 Person)

Sie benötigen für die Aufgabe eine aktuelle Python Installation mit mindestens Version 3.10 und Jupyter Notebook. Empfohlen wird Anaconda.

3) Aufgabenstellung

Schreiben Sie ein Jupyter Notebook, welches Daten aus mehreren Quellen einliest, aufarbeitet und eine Rangliste erstellt, bei welchem jede Gemeinde einen **Score zwischen 0 und 100** erhält. Beachten Sie, dass die Daten für einen fairen Vergleich von sämtlichen Gemeinden des Kantons berechnet werden müssen. Es ist aber erlaubt, für einzelne Features / Kriterien die Daten zu «cherry picken» - beispielsweise können Sie kategorisch allen einwohnerreichsten Gemeinden des Kantons einen Wert von 0 geben.

Verwenden Sie gesamthaft **mindestens 3 verschiedene Quellen**. Die Kriterien (Minimum 3) können frei gewählt und gewichtet werden und dürfen auch von einer anderen Aggregationsstufe wie z.B. Bezirk sein. Seien Sie kreativ, so dass ein unerwarteter Sieger entsteht und zwei (oder mehr) der Nachbarsgemeinden möglichst schlecht abschneiden. Der Einfachheit halber ist es erlaubt, Nachbarsgemeinde anstelle geographisch auch alphabetisch zu definieren (Die Nachbarsgemeinden von Windisch wären somit entweder Wiliberg und Wittnau, oder z.B. Brugg und Mülligen).

Erstellen Sie einen **kurzen Bericht** (ca. 1-2 Seiten A4) welcher zeigt, wie genial ihre gewählte Gemeinde ist. Der Bericht muss **mindestens 2 Visualisierungen** enthalten und soll die Daten möglichst irreführend präsentieren, so dass die Gemeinde so gut dasteht, wie es möglich ist.

4) Abgabe

Packen Sie Ihre Daten in ein .zip File mit dem Namen «dawr_1_vorname_nachname.zip». Dieses zip File soll enthalten: Ein Pdf mit dem Bericht (1 bis 2 Seiten), ein Jupyter Notebook mit dem Code sowie benötigte Files mit Daten falls nötig. Maximum 5 MB, wenn die Daten grösser sind, können Sie auch einen Link im zip oder im Notebook (NICHT im Mail) hinterlegen. Das Jupyter Notebook muss vollständig ausführbar sein. Verwenden Sie keine absoluten Pfade im Code.

Senden Sie das .zip bis am 23. März, 23:59:59 an lucas.broennimann@fhnw.ch

Code, welcher nicht von Ihnen selbst ist, muss entsprechend markiert sein. Plagiate werden mit der Note 1 bewertet. Nichteinhalten der Termine und Abgabebedingungen wird mit mindestens einer Note Abzug bestraft.