**Pomysł 1**

**Dataset road safety by WHO:**

<https://apps.who.int/gho/data/node.main.A989?lang=en>

Do przeanalizowania wpływ rozpowszechnienia problemów alkoholowych i nie tylko, przepisów ruchu drogowego, itp na liczbę ofiar wypadków drogowych i strukturę (ze względu na sposób w jaki się przemieszczali w chwili wypadku).

Datasety medyczne: <https://archive.ics.uci.edu/ml/datasets.php?format=mat&task=cla&att=num&area=life&numAtt=&numIns=greater1000&type=mvar&sort=dateUp&view=list>

mapka:

<https://extranet.who.int/roadsafety/death-on-the-roads/#alcohol/limit_novice>

**Pomysł 2:**

[Air pollution](https://www.epa.gov/outdoor-air-quality-data/download-daily-data) + zachorowania na astmę

Na stronie WHO, są dane odnośnie tego, ale generalnie tamte dane są często niekompletne i dotyczą tylko jakis wybranych typów i rodzajów zanieczyszczeń. Podobnie jest z resztą z tym ruchem drogowym.

**Pomysł 3:**

Mental illness + social media:

Mental illness w US dataset: [Link](https://www.datafiles.samhsa.gov/dataset/mental-health-client-level-data-2020-mh-cld-2020-ds0001) (od 2013-2020)

Co do social media usage to ciężko coś znaleźć ale jest raport z którego można zrobić jako taki dataset. [Link](https://www.pewresearch.org/internet/fact-sheet/social-media/)

Ale można stworzyć dataset do social mediów na podstawie trendów w google search: [Link](https://trends.google.com/trends/explore?date=all&geo=US&q=facebook)

Całkiem fajny pomysł, tylko problem jest taki, że tu nie będziemy mieli żadnej ścisłego połączenia pomiędzy mediami społecznościowymi, a zachorowaniami (mamy dane ogólne statystyczne o jednym i drugim, ale nie mamy danych typu osoba A o takich i takich cechach używała medium X => jest chora). To może być utrudnienie. Nie wiem co o tym myślicie?

**Pomysł 4:**  
ja miałem dwa pomysły, jeden to właśnie dokładnie o mental illness + social media, a drugi o **otyłości + fast foody**  
otyłość: <https://www.cdc.gov/obesity/data/adult.html>

fast foody: <https://data.world/datafiniti/fast-food-restaurants-across-america>

<https://www.kaggle.com/datasets/datafiniti/fast-food-restaurants>

Bardziej ogólne dane odnośnie otyłości (myślę, że dość proste do przeanalizowania):

<https://archive.ics.uci.edu/ml/datasets/Estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition+>

Można na przykład przeanalizować, jakie czynniki wpływają najmocniej na wagę osób i na ich otyłość oraz stworzyć model umożliwiający oszacowanie prawdopodobieństwa wystąpienia otyłości (prawdopodobną wagę) w zależności od podanych parametrów. Baza ma dużo rekordów, ale poszczególne wpisy mają tylko kilkanaście parametrów mających realny sens.

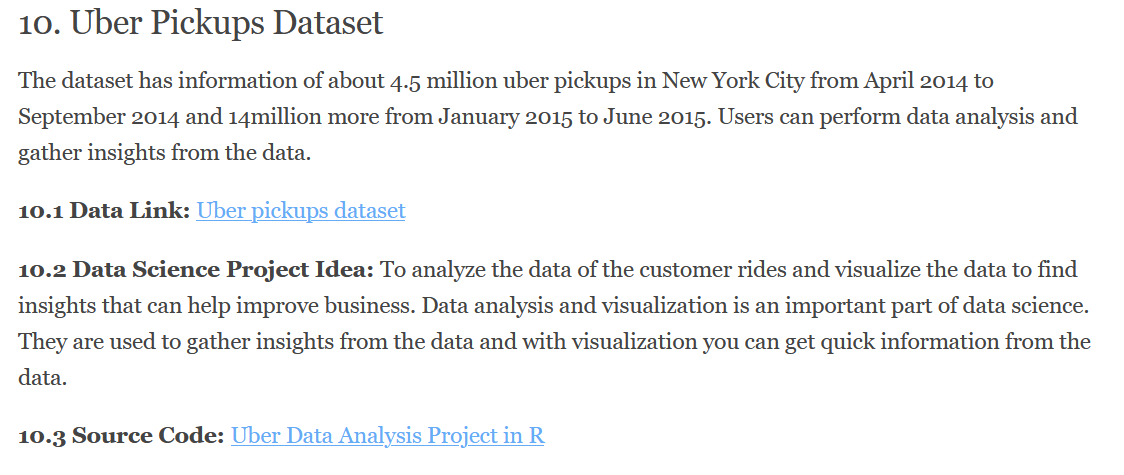
Można by te dwie rzeczy spróbować skleić w jedno.

**Pomysł 5:**

Analiza i klasyfikacja rytmu serca na podstawie danych zebranych w badaniu (też wygląda na nie nadmiernie trudne)(z datasetów medycznych). Tylko chyba nic poza analizą typu dane -> diagnoza nie da się z tego wyciągnąć. https://archive.ics.uci.edu/ml/datasets/Cardiotocography

**Pomysł 6:**

**https://data-flair.training/blogs/machine-learning-datasets/**



Generalnie ta strona zawiera fajne przykładowe tematy z podlinkowanymi datasetami od razu.

Proponuje niech każdy się zastanowi nad tym co się mu podoba i jutro do wieczora/niedziela rano napisze swoje opinie i spróbujemy ustalić jaki temat wybieramy.