Project Proposals

1152502 Jongho Park

1181873 Dulan

1164216 Marco Altieri

1171273 Aidan Prescott

Answer the following key questions:

•What is the research question?

Countless people worry about the environmental impact (contamination) of their society.

However, some people and businesses argue that pollution is not significant due to

environmental-conscious behavior and factory pollution. Pollution affecting people is no new

news, and although this is common knowledge, many people don’t seem to understand the

gravity of the situation By analysing how pollution affects the death of people, we hope to

pinpoint the severity of pollutants affecting the literal lives of people. We decided to look at the

impact of social pollution on people's health within different regions in Victoria.

How do various pollutants increase the amount of deaths which are avoidable in regions

throughout Victoria? (liveability and health)

•Why is it worth tackling?

- We can find the relation of pollution in the environment and diseases that cause death.

We can identify the relationship between the environment and people's health, including

pollution and lifestyle, identify factors that can improve people's health, promote people's health,

and alert people to factors that can cause diseases . This is especially important for individuals

who lived in the same area for extended periods of time, exposing themselves to the same

pollutants, people that settle are often elderly who are even more susceptible.

•What datasets can you use to solve it?

- 1. Avoidable death by suburb, extracted through an API.

- 2. Pollutions by suburb

Datasets of Avoidable death by suburb, Pollutions by suburb, Building plan by suburb from

Victoria government and Melbourne data-centre will be used for our purpose.

•What wrangling & analysis methods will you apply?

-Wrangling will be done via web scraping as the data is on an online api, we will also have

several CSVs on hand which will be processed via pandas

We look at each category of deaths and pollution in an attempt to find groups which relate can

be captured for further analysis. .

- And then we graph and compare and try to find patterns.

The dataset will be presented through graphs for comparing and showing the relationship

between unavoidable diseases and environmental factors.

•What are the expected outputs?

- We assume that there are some relations between certain diseases and pollution, which then

causes such deaths. This is because pollutants can contaminate food, water and air which upon

entering the body can cause health complications. In the long run, for people residing in the

area, they may develop illnesses, leading to death.

•What are the challenges and risks?

A large majority of the deaths are caused by diseases, specifically chronic such as cancer and

heart problems. These chronic diseases may be acquired from a previous location, in a different

environment hence the pollutant factors might be different. When analysing the data we have to

keep into account that the pollution does not directly cause the given death

Another challenge is that the investigation focuses on the cause of death rather than the

diagnosis, hence it is possible that pollution may have a bigger impact on the health hazard but

the fact that we are investigating deaths might overshadow this aspect. Although we should be

able to see a pattern for the death itself as if a certain area has more diagnosis for a certain

condition, then it is likely that it is a more common cause of death