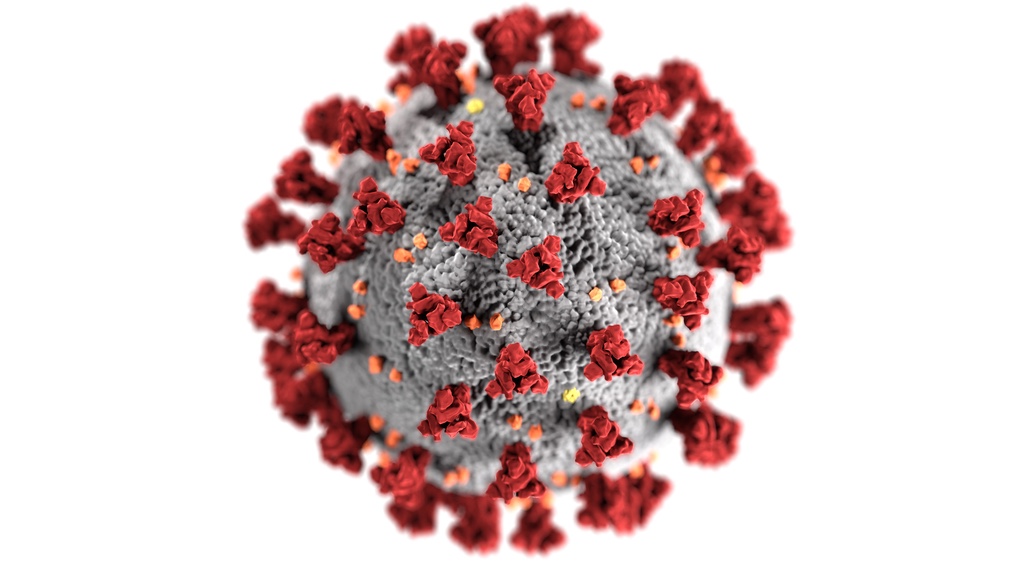
* [Overview](#gjdgxs)
* [Introduction](#30j0zll)
* [Assignment](#1fob9te)
* [Teacher's Solution](#3znysh7)

# CSE 450 - Machine Learning & Data Mining

## Exploration 05: Covid-19

 [*Image by CDC on Unsplash*](https://unsplash.com/photos/k0KRNtqcjfw)

## Overview

In this data exploration, you will:

* Explore Naive Bayesian Classification
* Learn to handle missing data
* Learn to encode textual data for machine learning
* Be introduced to SciKit Learn

Throughout this exploration, when you're asked to use a new function or library, we'll usually provide a link to that function's documentation, or a tutorial related to it.

## Introduction

As with our previous data explorations, this assignment uses [Google Colab](http://colab.research.google.com). For more information on using Google Colab, including how to submit assignments with it, please see the information in [Data Exploration 01](http://docs.google.com/exploration-01.html)

## Assignment

The data analyst firm you work for has been contracted by a local Health Authority to conduct a risk assessment of Covid-19.

Your research supervisor has asked you to use the latest CDC case data to preform a Bayesian analysis to calculate the probability of death from Covid-19 among certain demographics.

Click on the Open In Colab button below to open a Google Colab notebook with the template for this assignment. Once you've completed the assignment, don't forget to take the corresponding quiz in Canvas.



## Teacher's Solution

Once you have absolutely exhausted all of your best efforts in solving the data exploration problems, and you are stuck on where to go next, you can [view the teacher's solution here](https://github.com/byui-cse/cse450-course/blob/master/notebooks/Exploration_05_Solved.ipynb).