P1:

* What is my project
  + Analysis and data mining for crime patterns in an FBI data set from 1979-2014
* Why is data mining important for crime data? (reference: H. Chen, W. Chung, J.J. Xu, G. Wang, Y. Qin, M Chau, 2004, 50-56)
  + Mention those authors and their expertise
* What are different ways data mining can farm out effective patterns in crime data? (reference: H. Chen, W. Chung, J.J. Xu, G. Wang, Y. Qin, M Chau, 2004, 50-56)
* Mention the uniqueness of my project compared to that of the previously studied other projects (The ranking system, projecting data points onto more favorable or less favorable crime patterns over time).

P2:

* What are the different techniques used to solve similar types of problems? (reference: Krishnamurthy, R., Kumar S. J, 2012, 119-120)
  + Mention the 3 different types of algorithms that are most popular for this type of problem.
* What is generally considered a good approach, or attack to this type of project? (reference: Krishnamurthy, R., Kumar S. J, 2012, 118)
  + Mention the image and what is general considered a good practice approach to this type of problem.
* What results have others found from using the K-means clustering algorithm on the data? (reference: Nath, S. 2006, 18-22)
  + Explain some of the findings, short comings, and successes from this process. Also make mention again the distinction between their project and mine own.

P3:

* The algorithm I have chosen to use, and my justification (K-means).
  + Fast, gives a good initial idea of the behavior of the data, time constraints, ect …
  + If more time permitted, and I had better hardware to run my algorithm on, I would consider doing it with the EM clustering to compare the results and findings from the two.
* The approach I have chosen to take on the project.
  + Data cleaning, data integration, data transformation, how I chose to not do dimensionality reduction because of time (if time permits, I will do it to experiment with the results and findings and how they differ), then applying the K-means algorithm on a variety of combinations of the transformed data (including the population attribute and excluding it, centering the data and leaving it normal, I will apply on each combination from K=5,..30 by increments of 5.

Works Cited

Chen, H., Chung, W., Xu, J., Wang, G., Qin, Y., Chau M. Crime Data Mining: A General Framework and Some Examples (2004), 50-56 Aug. 2004, IEEE Computer Society April 2004. doi: [10.1109/MC.2004.1297301](https://doi.org/10.1109/MC.2004.1297301).

Nath, S. Crime Pattern Detection Using Data Mining, 18-22 Dec. 2006. IEEE. doi: [10.1109/WI-IATW.2006.55](https://doi.org/10.1109/WI-IATW.2006.55).

Krishnamurthy, R., Kumar S. J., Survey of Data Mining Techniques on Crime Data Analysis, Volume: 01, Issue: 02, 117-120, December 2012 International Journal of Data Mining Techniques and Application, ISSN: 2278-4219.