DRAFT Methodology:

\*\* Fix Data join code to remove the x & y occurences of “incident type”, then update EDA report. Will be linking github repo to my submission so want this to be cleaned up first. Don’t need to do any encoding because that only happens after split. I’m going to encode Activity Type and Response Type with the join (instead of creating lists to deal with duplicate rows for the incident numbers). That way there’s no dealing with lists later. Confirm with Professor Abdou that it’s fine that I did that.

As part of methodology, look up to see who else used this dataset before and what techniques and tools did they use. Check the efficiency of their classifiers. Check conclusions/recommendations that Parks Canada may have developed based on this data.

Should I group latitude/longitude values into X bins of areas within each park 🡪 group based on closeness so all latitude/longitude values within (for example) 5 km of each other (of centroid) are grouped together and then can be classified as high risk, med risk, or low risk areas based on number of incidents or specific types of incidents. I think this is outside the scope of my project 🡪 I would likely need to add more data so that I can capture the entire parks area and not just space where incidents happened. This project is more about predicting causes of certain incidents and or which incidents occur most in which parks, etc. I don’t need to do everything here.

3.

Human-Wildlife Coexistence: Recommendations for Improving Human-Wildlife Coexistence in the Bow Valley: <https://banff.ca/DocumentCenter/View/5520/Human-Wildlife-Coexistence-Bow-Valley-Report>

This report discusses current issues/trends with human and wildlife coexistence and plans to address it. The area of Bow Valley includes Banff National Park and Parks Canada technical experts were involved in the working group who put together the report. That said, there is no indication the that same dataset was used in the creation of this report and this report does not discuss the data source directly or how it was analyzed – only information regarding data is: “Data resources were compiled from the following agencies: • Parks Canada; • Government of Alberta; • Town of Banff; • Town of Canmore; and • WildSmart.”

This paper could be useful to refer to when making recommendations post analysis.

4.

Wildlife-Human Interactions in National Parks in Canada and the USA, Dr. Alistair J. Bath, Memorial University of Newfoundland Dr. Jody W. Enck, Cornell University, SOCIAL SCIENCE RESEARCH REVIEW: <http://npshistory.com/publications/wildlife/ssrr-v4n1.pdf>

This paper could be useful to refer to when making recommendations post analysis.

Quote:

9.

10.

I’m not super interested in this one – they focus a lot on the discrepancy between conflicts being reported for deprecation of livestock based on an animal when in fact the animal was not involved. Not super relevant to my project.

# A Framework for Estimating Human-Wildlife Conflict Probabilities Conditional on Species Occupancy

<https://www.frontiersin.org/articles/10.3389/fcosc.2021.679028/full>

“Managing human-wildlife conflicts (HWCs) is an important conservation objective for the many terrestrial landscapes dominated by humans. Forecasting where future conflicts are likely to occur and assessing risks to lives and livelihoods posed by wildlife are central to informing HWC management strategies.”

“We present a Bayesian hierarchical modeling framework that integrates conflict reporting data and species distribution data, thus allowing the estimation of the probability that conflicts with a species are reported from a site, conditional on the species being present. In doing so, our model corrects for both false-positive and false-negative conflict reporting errors. ”