

## M8 Lab 1C – Text Analysis with Python and NLTK

As biking becomes increasingly popular for both leisure and transportation, the issue of bike safety has never been more critical. Leveraging the power of Natural Language Processing (NLP) through Python and the toolkit, we analyzed textual descriptions of biking injuries. This analysis not only enumerates the most common types of injuries but also sheds light on potential focal points for improving bike safety.

Our text analysis focused on several body parts that were frequently mentioned in the injury descriptions we were able to extract. We were able to find out that biking injuries most frequently occur in the areas of the knees, shoulders, and elbows. This data is invaluable for safety gear manufacturers, suggesting a market need for high-quality protective gear targeting these body parts.

The word frequency analysis revealed a notable number of injuries described as occurring on the left and right sides of the body. While these terms may primarily indicate the side on which the accident happened, they still provide valuable information that could inform cyclist training programs. Such programs might include targeted educational efforts focusing on safer turning techniques or enhanced situational awareness at intersections and busy roads.

For instance, I have outlined several initiatives that could enhance bike safety based on the findings of this analysis.

- **Protective Gear Development:** The analysis highlights the vulnerability of certain body parts like knees, shoulders, and elbows. Manufacturers can utilize this information to create more specialized gear. For example, knee pads with extra cushioning or braces that provide additional lateral support could be developed. Similarly, shoulder guards designed for bikers could be introduced, prioritizing mobility as well as protection.
- **Infrastructure Improvements:** Knowing which types of injuries are most common can help city planners design better bike lanes or cycling paths. For instance, if shoulder injuries are frequent due to collisions at intersections, planners could implement more visible signage or even separate cyclist-only traffic lights to reduce the risk.
- **Educational Programs:** Data on common injuries can guide the curriculum of bike safety courses. If knee injuries are prevalent, these courses could include specific exercises to improve knee strength or tutorials on how to fall in a way that minimizes knee impact.
- **Policy Implications:** The prevalence of specific injuries, such as those to the knees and elbows, could serve as a catalyst for policy changes at both the local and national levels. Legislation requiring mandatory protective gear like knee and elbow pads in areas

identified as high-risk could be enacted to mitigate the severity and frequency of these injuries. This is also connected to the infrastructure improvements.

In summary, our textual analysis has provided key insights into the types of injuries most often associated with biking accidents. By pinpointing the body parts most vulnerable to injury, we can inform the design of targeted safety gear and educational programs. Additionally, these findings can be invaluable for medical professionals who are often the first line of treatment after an accident. As biking continues to grow in popularity, the application of data analysis in enhancing bike safety cannot be overstated. The insights gained through this research lay a robust foundation for multi-faceted approaches to significantly reduce biking-related injuries.