**Ski Resort Reviews and Resort Statistics**

**Python – Andrew Boral**

This project explores two datasets to find out more about ski centers in the United States.

**Ski Center Reviews:**

Data Source: <https://www.kaggle.com/datasets/fredkellner/onthesnow-ski-area-reviews>

Data File: OnTheSnow\_SkiAreaReviews.csv

My Python Code to Analyze Source:

1. FinalProject NLTK By Review.ipynb
2. FinalProject NLTK By Ski Center.ipynb

Description:

The first dataset is composed over 16,000 text reviews. Natural Language Processing tools were used to analyze the reviews. The data was preprocessed to remove capitalization, remove punction and numbers, and remove common English stop words. Next, TF-IDF analysis was performed to reduce the weight of stop words common to this corpus.

The python libraries utilized to complete the analysis were NLTK and Gensim.

Results:

**Ski Center Statistics:**

Data Source: <https://www.kaggle.com/code/beaubellamy/ski-resorts>

Data File: skiResort.xlsx

My Python Code to Analyze Source:

1. FinalProject-SkiResortStats.ipynb

Description:

Results: