

Topic 2: Popularity of Munros

1 Overview

In this research I will be discussing the factors affecting the popularity of Munros.

2 Introduction

Context and motivation What is the area of this data science study, and why is it interesting to investigate?

Objectives What questions are you setting out to answer?

The main question I will be exploring is what factors make a munro popular? This includes answering questions: What factors have the most influence in the popularity of munros? How do the natural relief and cities' location affect the popularity?

3 Data

At the beginning of the research I have used data provided by <https://www.walkhighlands.co.uk/munros/>. However, after doing some work on it, I have realised that the data in <https://www.walkhighlands.co.uk/munros/> dataset is not enough for me to draw any certain conclusions and I have also decided to use data provided by <http://www.hills-database.co.uk/downloads>.

Data provenance Who created the dataset(s)? How you have obtained it (e.g., file or web scraping), and do the T&Cs allow you to use obtain the data for the project?

The datasets I have used are provided by <https://www.walkhighlands.co.uk/munros/> and <http://www.hills-database.co.uk/downloads>. Walkhighlands provided data in tables on the website which I have then converted in csv format files. Hills-database provided a csv file. During the research I have also done my research on population and location(in coordinates) of biggest cities in Scotland stored in cities data. This data was provided by <https://www.latlong.net/>.

Data description Description of the data, e.g. variables in each table, number of records.

In the first version of the data set which I have created from Walkhighlands's data there were 282 entries of munros. It contained information about altitude, region, walk reports and ascents number and rating. In the second version, the number of entries reduced to 218 due to the differences in the munros data in Hills-database (Figure 1). New columns with the location of munros and distances from the major cities were added (Figure 2).

	Mountain	Altitude m	Region	Ascents	Popularity	Grid Ref	GridRefXY	Longitude	Latitude
0	Ben Nevis	1345	Fort William	17927	2	NN166712	NN 16666 71288	56.796907	-5.005963
1	Braeriach	1296	Cairngorms	6044	2	NN953999	NN 95327 99915	57.078312	-3.728821
2	Cairn Toul	1291	Cairngorms	5127	1	NN963972	NN 96328 97229	57.054420	-3.711193
3	Sgor an Lochain Uaine	1258	Cairngorms	4837	1	NN954976	NN 95422 97693	57.058382	-3.726326
4	Cairn Gorm	1245	Cairngorms	12271	2	NJ005040	NJ 00516 04065	57.116727	-3.644851
...
213	Ruadh Stac Mor	918	Ullapool	2133	0	NH018756	NH 01853 75655	57.726806	-5.332417
214	Carn Aosda	917	Cairngorms	11206	2	NO133791	NO 13397 79164	56.895715	-3.423520
215	Beinn a' Chleibh	916	Argyll	6698	2	NN250256	NN 25054 25606	56.390246	-4.837547
216	Beinn Teallach	915	Fort William	3754	1	NN361859	NN 36136 85973	56.935947	-4.696405
217	Ben Vane	915	Loch Lomond	10491	2	NN277098	NN 27753 09844	56.249796	-4.783462

218 rows × 9 columns

Figure 1: Final version of the data describing munros after merging the datasets from <https://www.walkhighlands.co.uk/munros/> and <http://www.hills-database.co.uk/downloads>, cleaning the data and converting grid reference into coordinates.

Data processing How you have processed the dataset, e.g., cleaning, removing missing values, joining tables.

The first step was to merge data from Walkhighlands. I used pandas to create a data frame after loading the csv files I have created earlier. This step didn't need much cleaning as all datasets provided by Walkhighlands had the same number of munro entries and same naming. The second step was to add data from Hills-database. Most of the information was irrelevant, and I have only left column that contained Grid Reference. Since I preferred to work with coordinates, I have converted this data using OSGridConverter library. After I have created a new dataframe which stored the coordinated of location of seven major Scotland cities such as Edinburgh, Glasgow and Perth. After obtaining latitude and longitude data for munroes and cities, I have calculated the distance from each munro to each city using Haversine distance (Figure ??) [1].

4 Exploration and analysis

	Cities	Population	Longitude	Latitude
0	Edinburgh	482005	55.9533	-3.1883
1	Glasgow	598830	55.8642	-4.2518
2	Dundee	148300	56.4620	-2.9707
3	Aberdeen	207932	57.1497	-2.0943
4	Inverness	46870	57.4778	-4.2247
5	Stirling	94330	56.1165	-3.9369
6	Perth	47430	56.3950	-3.4308
7	St Andrews	17580	56.3398	-2.7967

Figure 2: Data of the chosen cities in Scotland

Notes: To measure the popularity of Munroes I have considered the total number of ascends. Rating is not a suitable measure of popularity as it is very subjective to each individual and not all people leave ratings, which makes it unreliable. For better understanding, I have also created three categories of popularity: Unpopular(0) with the number of ascended below 40% of the maximum number of

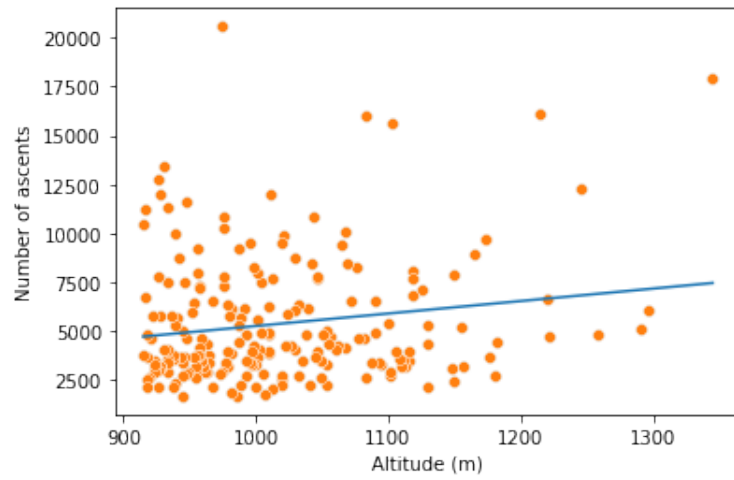


Figure 3: Initial attempt at plotting frequency of ascents and altitude.

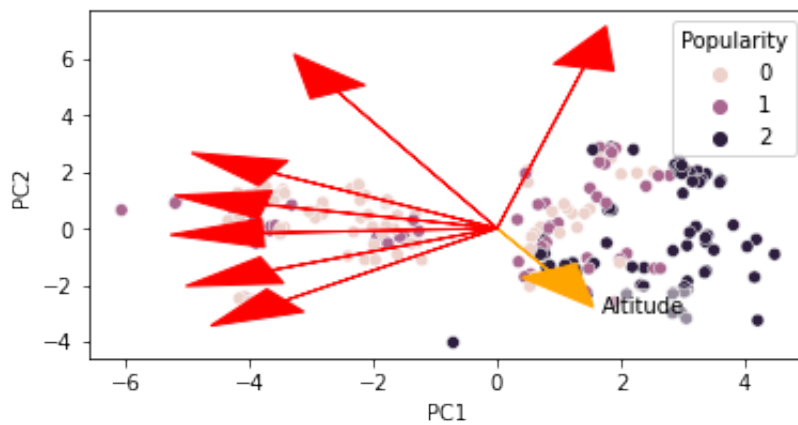


Figure 4: Influence of original variables on the PC scores. Notation: orange arrow representing altitude, red arrows represent distances from cities, from bottom to top: Edinburgh, Glasgow, Dundee, Aberdeen, Inverness, Stirling, Perth and St Andrews. Not included to keep the graph clear.

ascends, Medium popular(1) - above 40% and below 70%, and Popular(2) - above 70%. Process: At the early stages of the research, I have tried different approaches to manipulating the data. I started with plotting the pair plot on the first version of the data to see any obvious patterns. Unfortunately, there weren't any obvious patterns. Plotting a graph between the altitude of munros and the number of ascends, as was suggested in the description of the coursework, also didn't give any convincing arguments, with the correlation coefficient being only 0.13668 (Figure 3). This, however, suggested that height affect popularity, but it is not the only factor. After examining the map of munros from <https://www.walkhighlands.co.uk/Forum/memberlist.php?mode=viewmap>, I have decided to create another dataset with the distances from munros to the top largest cities in Scotland, such as Edinburgh, Glasgow and Perth. To check this I have conducted PCA on the data (Figure 4). This PCA describes 83% of the variance: PC1 explains 63% and PC2 explains 20%. PCA has also revealed multiple clusters. Interestingly, clusters did coincide with the established popularity ratings, even though the separation between them is not very clear. The reason behind this is other possible munro characteristics or possible details of the relief, which I am not focusing on in this report. This suggests that the popularity of munros does indeed depend on the distances from the cities.

I have decided to look more into the correlation between distances from the cities and number of ascends. Comparing different cities revealed an interesting pattern, which appears for most of the cities (Figure 5). I believe that there should be an exponential relation between these two variables, however, since we haven't studied it during the course I have decided to show how a linear regression would not fit the data perfectly. The exception is Inverness, which could be due to the close locations of the similar munros, therefore people don't have any particular preferences when choosing which one to climb.

And finally, I have explored the idea of the popularity of munro depending on where it is located in Scotland geographically (Figure 6). This describes additionally 15% of the variance in the data. I have also put the names of top 5% most popular munros on this figure.

5 Discussion and conclusions

Summary of findings Overall, there are multiple findings I have made in this research. Popularity of munros depend on how far it is from the cities, which can be interpreted that most of the people would choose a munro close to where they live/stay to visit. Popularity also depends on the height, higher munro usually tend to be popular, perhaps due to the fact that people want to challenge themselves and for the beautiful views, however that doesn't work for very high munros. The exceptions is Ben Nevis, as it is the highest munro but is the most popular one. But overall a lot more ascents are on the lower munros. The most popular munros are located to the South, South-East, while munro to the North are a lot less popular, perhaps due to the difficulties of travelling there.

Evaluation of own work: strengths and limitations One of the strongest aspects of this work is the amount of research done to create the data set, as well as aspects such as how to find the distance between two points using their coordinates taking in account curvature of Earth. I also think that doing PCA on the data was definitely the right strategy and I am quite impressed with how much information it provided. As for the limitations, my knowledge of Geography. It believe that if I had better understanding of reliefs and munros I would be able to derive more conclusions. I would And something that was really hard for me personally was timing, I had a very busy period with multiple interviews and a spring week clashing at the same time with this deadline.

Comparison with any other related work Conclusion about Ben Nevis was supported by another dataset of the best munros [2].

Improvements and extensions As I have mentioned earlier I would improve the regression to fit the data in Figure 5 better. Another improvement is getting better at Latex, as I have never worked with

Number of ascents against the distance (km) from the major Scottish cities



Figure 5: Finding correlation between distances from the cities and number of ascends.

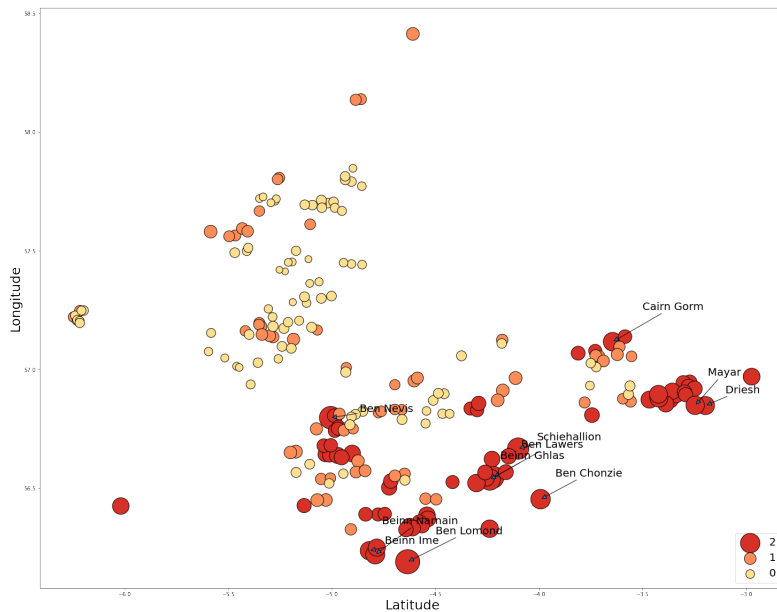


Figure 6: Popularity of munros depending on the georgaphical direction.

it before and it was quite tricky for me to start learning it from scratch. I tried my best to make this document look as described in the requirements. As for extensions, I would definitively look more closely to the influence of the landscape and relief and aspects like view from the munro that might attract people or difficulty level.

6 Bibliography

- [1] How to calculate the Euclidean distance between all points of latitude LONGITUDE PAIRS? (n.d.). Retrieved April 12, 2021, from <https://uk.mathworks.com/matlabcentral/answers/519457-how-to-calculate-the-euclidean-distance-beetwen-all-points-of-latitude-longitude-pairs> [2] 10 best MUNROS. (n.d.). Retrieved April 13, 2021, from <https://www.stevenfallon.co.uk/munros-best.html>