

The tourist and recreational offer along with services available for tourists in Poznań don't create strong, unique brand of city - one that would be recognizable worldwide and could be compared to Kraków's, Gdańsk's or Warszawa's. In the ranking of the top attractive cities published in 2018 by ARC, Poznań has been placed less attractive than Wrocław or Gdańsk. All of these cities were also much below Kraków's. We are sure that Poznań, a beautiful and friendly city, has a lot to offer for guests and could attract more tourists. Are you ready to help the city of Poznań to explain its touristic popularity?

HackRcity challenge

HackRcity aims to understand the touristic popularity of Polish cities, provides solution on **what impacts most the tourist-friendliness of Poznań** and finally, plans a schedule of a family going on holidays taking into account facts about cities and family's preferences. We want to leverage the extensive experience of Analyx in predictive modelling and optimization to help data science enthusiasts in addressing this social problem.

EXPLORE, PREDICT & EXPLAIN: Find out why tourists are coming to Poznań.

- How to predict the touristic popularity index (in particular cities marked as 'NA')?
- How to predict the number of tourists coming to Poznań?
- Which factors explain the touristic popularity and which explain the number of tourists?
- Do tourists want to visit museums, eat in a fancy restaurant, or stroll around a park?
- Are these factors the same as in Gdańsk, Kraków, Warszawa, or Łódź? How does Poznań look in comparison to other Polish cities?
- How do you see the future of tourism in Poznań?
- What needs to be changed in Poznań in order to attract more visitors, assuming that geographic location remains the same (we will not move oceans nor mountains)?
- Why other city of similar size (e.g., Wrocław) has more tourists than Poznań?
- Which actions are recommended for Poznań to go up in the touristic popularity ranking?

OPTIMIZE: Plan holidays of a family (see file with preferences) knowing what are the distances between cities. You need to start and finish in the same city and travel not more than 1500 km in total, visiting not more than 5 cities.

- Which cities will you recommend for the family to visit?
- What should be changed in Poznań to be on this holiday plan?
- What should change in your approach/assumptions to make Poznań appear in this plan?

PRESENT your results and impress us!

About data

At 10.15 each team receives access to the same data frame with some data from GUS (Główny Urząd Statystyczny), Wikipedia and distances between cities. Data covers 66 main Polish cities – all names of the cities will be hidden behind fruit names, except for Poznań (see folder “Dane”).

At 11.00, each team receives **7 HackRpoints** that can be exchanged for additional data (which **can but not have to be** useful for predicting a target). **NOTE:** if the team does not exchange some of the HackRpoints, they disappear at 18.00 (they are not summed up to the final score).

HackRpoints exchange rates

Additional data	HackRpoints
Dataset with 10 real city names (randomly selected)	2
Restaurants data from a leading travel portal	2
Attractions data from a leading travel portal	2
Shopping data from a leading travel portal	2
Accommodation data from a leading travel portal	2
Additional data from GUS	2
Data of the sentiment analysis from popular social network	3

Webscraping is allowed free of charge 😊

Some additional facts about data or challenge may be revealed during the event until 17:00. They will be made available to all the participants at the same time with the use of **Slack**.

Expected outcomes

19:00 Forecasts of the touristic popularity of cities marked as ‘NA’ shared in folder “Wyniki” as data frame called ‘1_output_df’ with 7 rows for cities which touristic popularity index is missing and 2 columns called ‘city’, ‘touristic_popularity’. **NOTE:** the values for column ‘city’ must be as they were originally given (6 fruits + Poznań).

19:00 All code shared in folder “Wyniki”

19:55 Slides for presentation (optionally with supplementary materials) shared in folder “Wyniki”

19:55 All code updated in folder “Wyniki”

20:00 Oral presentation of the approach and results

Evaluation criteria

	Evaluation criteria	Weight	Task
1	Accuracy of the solution calculated from the mean absolute error (MAE) of the forecasts for the touristic popularity of 7 indicated cities (marked as 'NA' in the dataset) of year 2018. NOTE: an error in prediction of Poznań's touristic popularity index will be multiplied by 2	20%	EXPLORE, PREDICT & EXPLAIN
2	Interpretability of the solution assessed as opinion of the jury experts about attractiveness of their explanations, conclusions and recommendations	20%	EXPLORE, PREDICT & EXPLAIN
3	Reproducibility of the solution assessed as opinion of the jury experts about potential future use of code	20%	ALL
4	Quality of solving holiday planning assessed as opinion of the jury experts about originality of the proposed approach	10%	OPTIMIZE
5	Advancement of research assessed as opinion of the jury experts about the variety of the methodologies and concepts used and the overall robustness of the solution	10%	ALL
6	Presentation of results assessed as opinion of the jury experts about the oral presentation of results and approach and the quality of the presented slides	10%	PRESENT
7	Usefulness of results assessed as opinion of an external expert, CEO from the Poznań Tourism Organization	10%	PRESENT