PIZZA VS. DÖNER

A REGIONAL STATISTICAL ANALYSIS ON THE PIZZA AND DÖNER RESTAURANTS DISTRIBUTION IN GERMANY.

INSIGHTS INTO SOCIO ECONOMIC FACTORS FOR OPENING A PIZZA OR DÖNER PLACE IN GERMANY

- Both Pizza and Döner Kebab have a high priority in the dietary plan of the average German.
- However, regions in Germany differ widely in terms of:
  - Population density
  - Available income
  - GDP
  - Size and inhabitants
  - Count of Pizza and Döner places per capita
  - Preference to either Döner or Pizza by the inhabitants
- This statistical analysis will work out key statistical factors that can assist on the decision whether to open a Döner or a Pizza restaurant and especially were.

DATA SOURCES

1. Geo data for all regions in Germany including their geometry coordinates for choropleth plotting in Json format: https://github.com/isellsoap/deutschlandGeoJSON
   - Name of the region
   - Border geometry coordinates
2. Statistical data from the German ministry for statistics destatis.de: https://www-genesis.destatis.de/genesis/online
   - Availible income per person
   - GDP per person
   - Region population
   - All data is from 2018
3. Restaurant location and information using the Foursquare API: https://developer.foursquare.com/developer/
   - „Search“ enpoint is used to determine exact type/group of the restaurant and the search result count.

DATA CLEANING AND PREPARATION

1. To merge the Geo data and the data from the german institute for statistics (destatis.de) a custom merge algorithm was required.
2. The merged dataframe was used to generate search coordinates and search radii for each region for the search endpoint of the Foursquare API.
3. The resulting features selected was the count for the pizza and döner places respecitvly.
4. From the now availible data, a few augmented fields field generated that will be plotted in the the next section.

CORRELATION ANALYSIS

More pizza restaurants correlates slightly with also more restaurants in a region.

More inhabitants correlates slightly with more restaurants in a region.

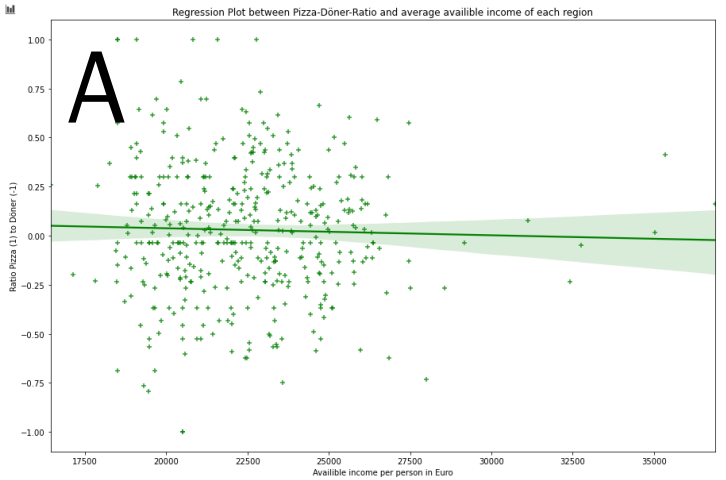
Other correlation are obvious (derived values).

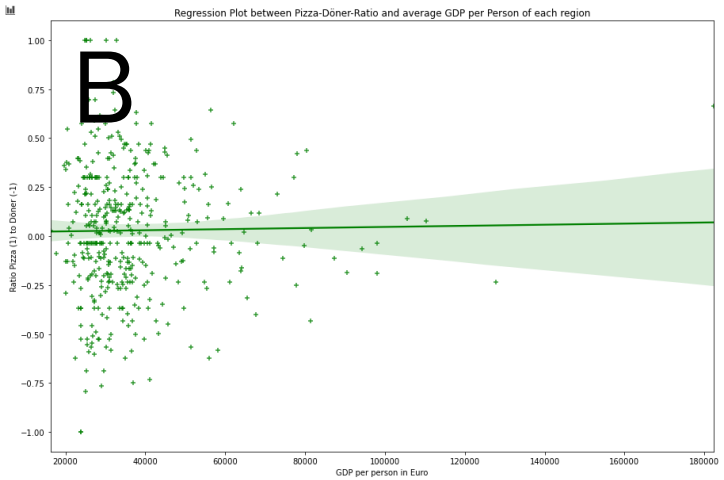
Correlation Analysis

	ve_pp	bip_pp	Einwohner	doener	pizza	ratio	density_100k
ve_pp	1.00	0.30	0.05	0.19	0.18	-0.03	0.19
bip_pp	0.30	1.00	0.09	0.19	0.23	0.02	0.14
Einwohner	0.05	0.09	1.00	0.46	0.51	-0.01	-0.17
doener	0.19	0.19	0.46	1.00	0.81	-0.31	0.51
pizza	0.18	0.23	0.51	0.81	1.00	0.14	0.45
ratio	-0.03	0.02	-0.01	-0.31	0.14	1.00	-0.12
density_100k	0.19	0.14	-0.17	0.51	0.45	-0.12	1.00

REGRESSION ANALYSIS

1. There is no over-all positive or negative relationship between the wealth of a region and the ratio of pizza to döner. (A,B) The long-lasting questions if Pizza or Döner is more classy hence can not be answered with this analysis :(
2. There is a significant positive correlation between the density of Pizza+Döner restaurants and the available income reaching for ~10 per 100k for poor regions to ~20 per 100k for rich regions (C). This implies that it is a good idea to choose a region that has high wealth (and ideally not too many restaurants already).

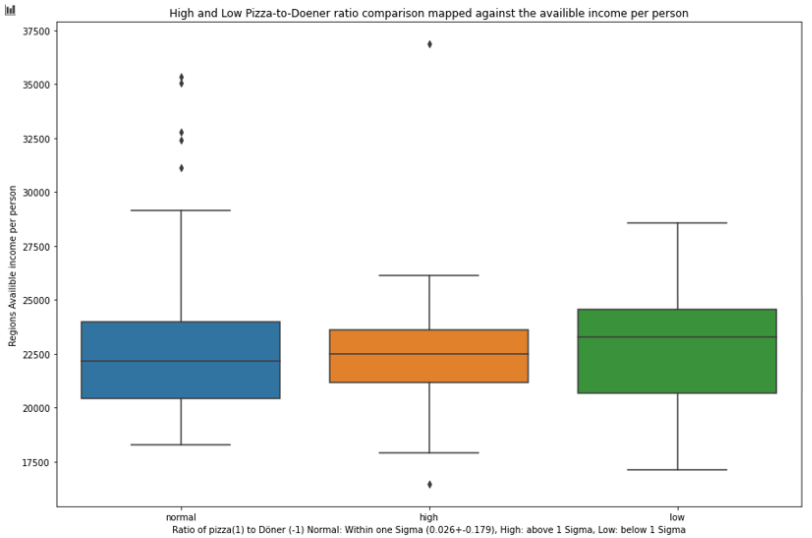




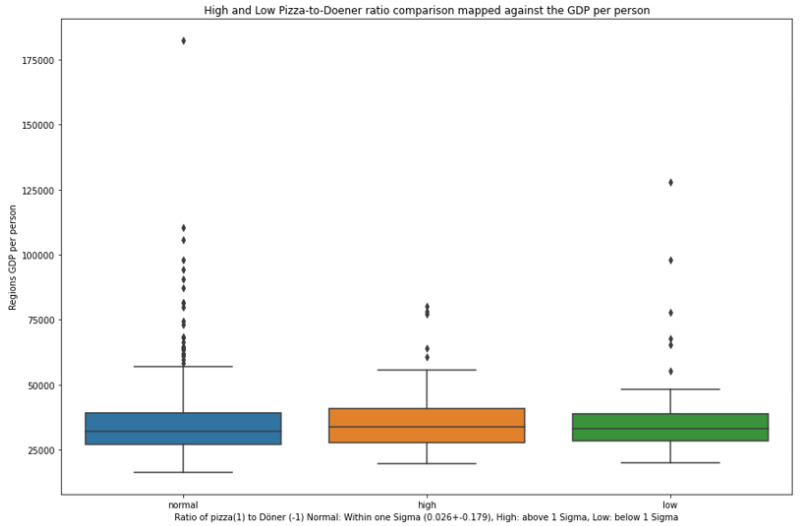
SPLIT GROUP ANALYSIS – IS PIZZA OR DÖNER THE MORE CLASSY?

There is no significant difference found in any of the plots again suggesting that Pizza and Döner are equally classy food.

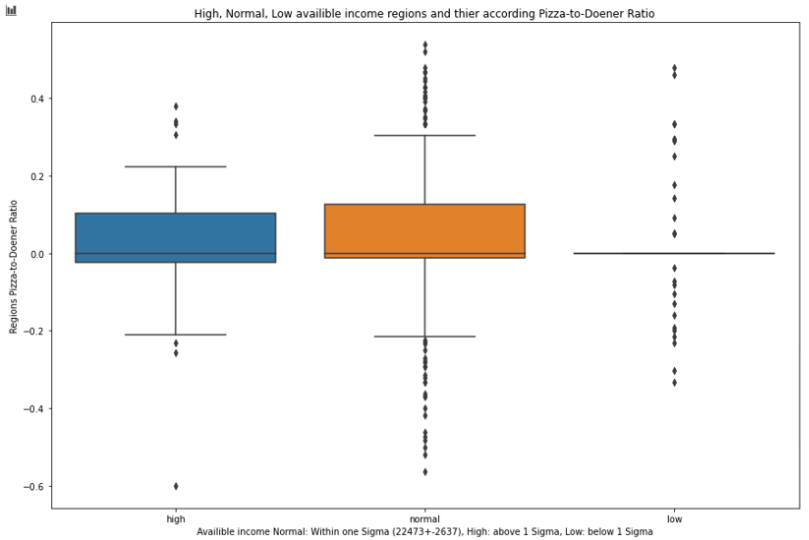
This analysis only looks at regions not at individuals. So, it can not ultimately answer this questions.



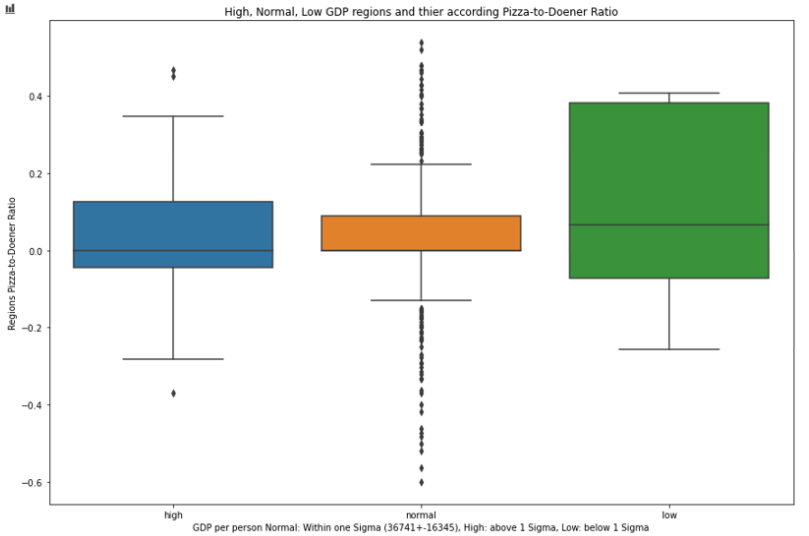
Tile. A High and Low Pizza-to-Doener ratio comparison mapped against the availible income per person.



Tile. B High and Low Pizza-to-Doener ratio comparison mapped against the GDP per person.



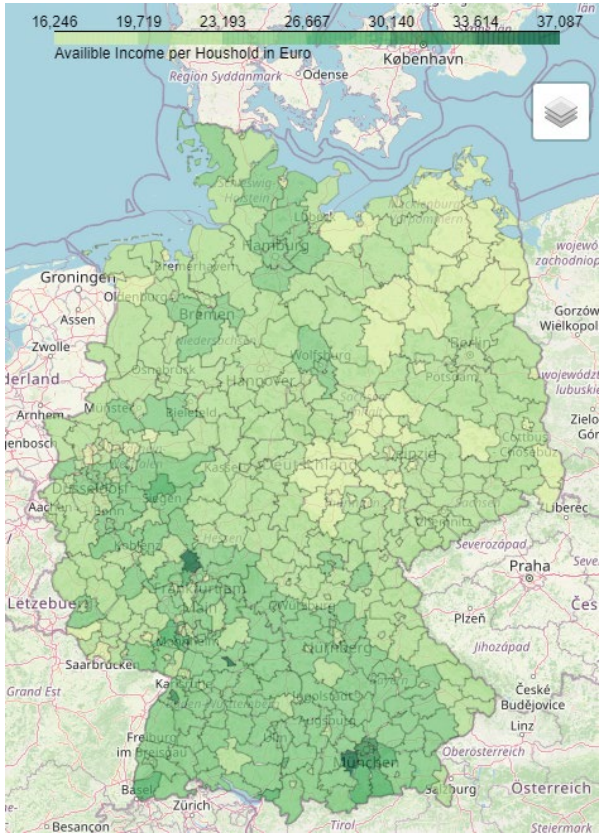
Tile. C High, Normal, Low availible income regions and thier according Pizza-to-Doener Ratio.



Tile. D High, Normal, Low GDP regions and their according Pizza-to-Doener Ratio.

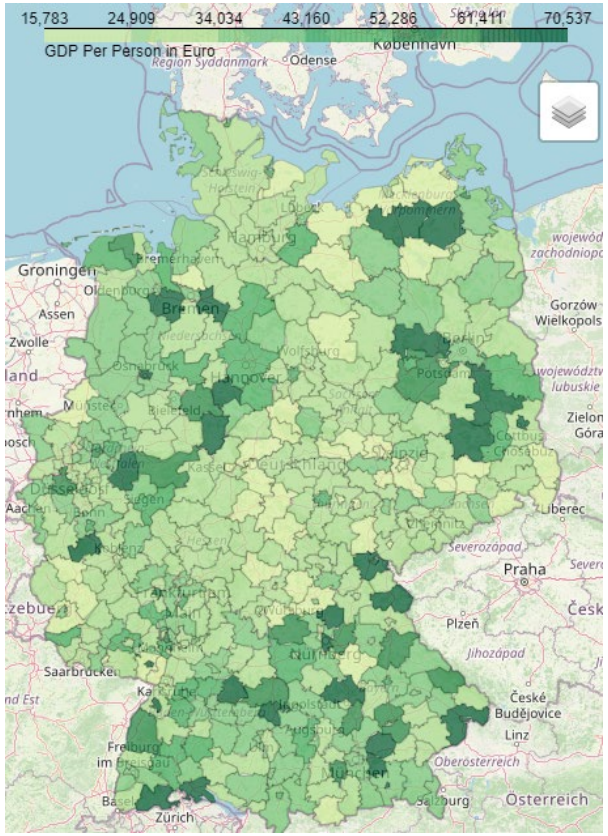
WHERE IS THE WEALTH AND THE PRODUCTIVITY CONCENTRATED?

Available income per household in Euro.



The south and the south-west are wealth regions, but exceptions are areas near major cities like Hamburg.

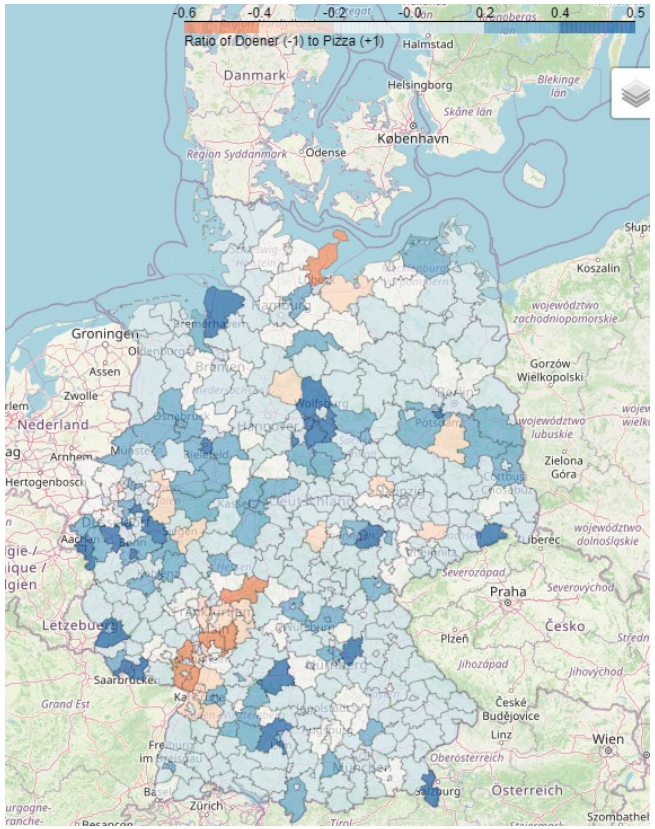
GDP per person in Euro



GDP is dominated by regions with large companies.

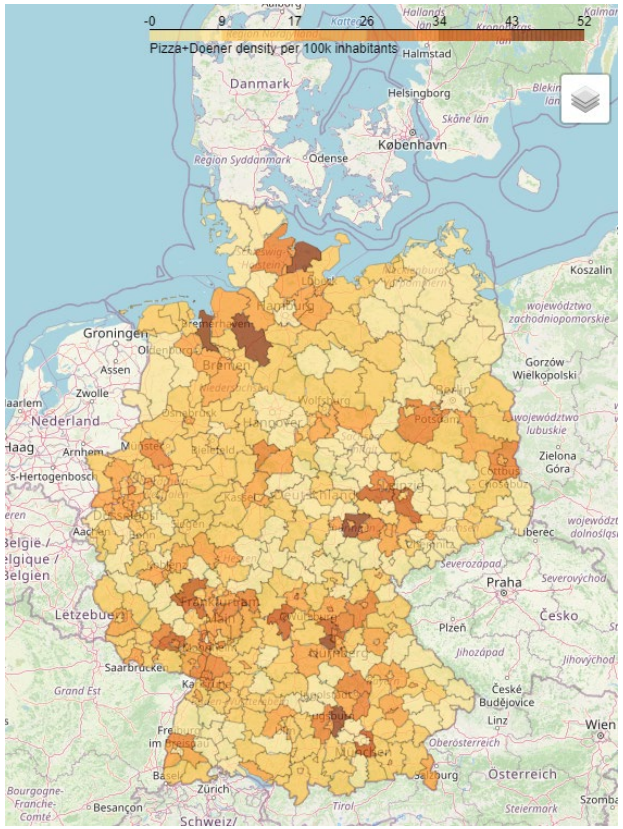
WHICH REGION PREFERS WHAT AND HOW FIERCE IS THE COMPETITION?

The ratio of döner (-1) to pizza (+1)



There are regions and clusters with clear dominance of either Pizza or Doener. It hence would make sense to choose the region accordingly and choose the richest one.

The Pizza+Döner density of each region per 100k inh.



The density differs widely, meaning it is a factor that should be respected in the final choice of the region.

RECOMMENDATIONS: BEST DÖNER REGIONS

The recommendation is to open a Döner in "Bad Dürkheim" region where people apparently crave the Döner Kebab and are loaded.

"Main-Kinzig-Kreis" would be a good Döner alternative, it is not quite as rich but they only have 1/3rd of the restaurant density there compared to "Bad Dürkheim".

Gebietseinheit	ve_pp	bip_pp	density_100k	availible_income_rel_to_avg	GDP_rel_to_avg
Bad Dürkheim, Landkreis	26863	22551	22.556391	4390	-14190
Main-Kinzig-Kreis	24607	36532	7.637232	2134	-209
Odenwaldkreis	22564	26668	25.773196	91	-10073
Bergstraße, Landkreis	24752	28488	19.330855	2279	-8253
Ostholstein, Kreis	23179	26919	13.432836	706	-9822
Germersheim, Landkreis	23173	37489	29.457364	700	748
Südliche Weinstraße, Landkreis	24538	27569	23.636364	2065	-9172
Groß-Gerau, Landkreis	22014	45716	17.582418	-459	8975
Miltenberg, Landkreis	23343	34398	23.255814	870	-2343
Calw, Landkreis	24437	29286	12.025316	1964	-7455

ve_pp: The availibe income per person.

bip_pp: The GDP per person.

density_100k: the restaurant (döner+pizza) density per 100k inhabitants.

Available_income_rel: the available income per household relative to the mean.

GDP_rel: The GDP relative to the mean.

RECOMMENDATIONS: BEST PIZZA REGIONS

"Bonn“ is a good bet where the people would sell their grandma for a good pizza; but they don't have to because they have plenty of cash and the productivity of the region is insanely good.

"Rhein-Erft-Kreis" would be the alternative for Pizza, due to an severe lack of good restaurants per 100k inhabitants and a high pizza preference and an still average income.

Gebietseinheit	ve_pp	bip_pp	density_100k	availible_income_rel_to_avg	GDP_rel_to_avg
Berchtesgadener Land, Landkreis	22547	33571	24.761905	74	-3170
Leverkusen, Kreisfreie Stadt	21756	51469	15.243902	-717	14728
Rhein-Erft-Kreis	22675	34641	4.904051	202	-2100
Jena, Kreisfreie Stadt	18919	44877	20.720721	-3554	8136
Bonn, Kreisfreie Stadt	23471	80288	13.803681	998	43547
Bielefeld, Kreisfreie Stadt	23037	41016	9.009009	564	4275
Wuppertal, Kreisfreie Stadt	21480	36429	12.711864	-993	-312
Saarbrücken, Regionalverband	19205	44904	19.090909	-3268	8163
Braunschweig, Kreisfreie Stadt	22612	78057	16.129032	139	41316
Alb-Donau-Kreis	23854	32093	18.461538	1381	-4648

ve_pp: The availibe income per person.

bip_pp: The GDP per person.

density_100k: the restaurant (döner+pizza) density per 100k inhabitants.

Available_income_rel: the available income per household relative to the mean.

GDP_rel: The GDP relative to the mean.

CONCLUSION

1. There is no regional correlation between wealth and a Döner or Pizza preference.
2. Density of restaurants to inhabitants differs widely but does not correlate with pizza or döner.
3. Hence there is not clear recommendation for a Pizza or Döner region.
4. There is a correlation between wealth and count of restaurants. Following this logic it is an important factor to start a restaurant in a wealthy area with a not yet high restaurant density.
5. Döner recommended areas for starting a business are Bad Dürkheim and Main-Kinzig-Kreis due to their heavy döner preference and good income with moderate restaurant density.
6. Pizza recommended areas are Bonn and Rhein-Erft-Kreis employing the same logic.