<u>Migration flows in Vienna</u>

from 2002 to 2016
M4

1 Motivation

The motivation for this project is the currentness and the timelessness of the topic. Austria always gave home to lots of people from everywhere around the globe. The collapse of the Soviet Union brought thousands of people to Austria, during the Balkan Wars about 100.000 people found home in Austria and the newest wave of refugees from West Asia and North Africa is not over yet. Besides all those people fleeing from war there are thousands who came in hope of a better economical future or academic chances.
A noteworthy amount of all those immigrants end in Vienna. This may be no surprise due to the meanwhile achieved multiculturalism and top quality of life. Vienna developed and still develops because of the people who are living there. For us, living in Vienna and knowing the city, it is very interesting exploring why one district looks like it does, with all its shops, markets and other idiosyncrasies. Beyond the borders of Vienna, people in Austria know that the 10th district, Favoriten, is "full" of foreigners but we want to take a closer look at all that.

2 Related Work

A similar problem is discussed in this paper[1] and they did come up with a really nice solution in our opinion, but such diagrams aren't possible with tableau. This group managed to create circular plots for the visualisation of migration flows. This paper gave us the idea to try and create sankey diagrams in tableau, because we really liked this type of visualization. Although Tableau is a really powerful tool as we thought at the beginning it does has its limits and several times we needed do some workarounds to create what we had imagined, inspired by [1]'s work.

Another website we looked at was [3]. It shows all world's countries and their immigrant and emigrant numbers. On the left side of the webpage you see countries and their emigrant-info, on the other side you see countries and their immigrant info. This is a really simple, but informative and clean way to show this one-sided data.

Our idea and how to actually create a sankey diagram we got from this website[2]. Sankey diagrams are really awesome to show actual flows between something and were perfect for our topic. Creating sankey diagrams is quite difficult in tableau, because you have to do a lot of calculations to draw some curves in tableau and sankey diagrams are not featured in

---

[1] Vienna Institute of Working Demography Working Papers 2/2014; Nikola Sander, Guy J.Abel, Ramon Bauer and Johannes Schmidt; Visualising Migration Flow Data with Circular Plots http://www.global-migration.info/VID%20WP%20Visualising%20Migration%20Flow%20Data%20with%20Circular%20Plots.pdf
[2] https://www.superdatascience.com/yt-tableau-custom-charts-series/
[3] http://peoplemov.in/

Tableau. Although it wasn't our idea how to create sankey diagrams in tableau, we had to adjust the technique so we could implement it to our project.

3 Approach

3.1 Project Proposal

Milestone 1 was about finding answers to general questions about the project. Such were the theme to discuss, type of project, target users, dataset, expected problems etc. The major task here was to find our theme. We invested time to compare themes and datasets we were able to find, there were several other interesting themes than our chosen one, but datasets were, when found at all, too small. Finally the team agreed on the following dataset, the reasons for the choice were mentioned above in *1 Motivation.*

3.2 Lo-Fi Prototyping

When starting with the project we had very little experience with Tableau and the dataset we found was pretty big and confusing at some points, but we had our ideas and so we sketched them during milestone 2. This was, more or less, the beginning phase of the project. To see what was possible with Tableau we made some quick visualizations to see what kind of charts are doable, so our prototypes would not be totally different to the actual implementation. During the course Visualization and Data Analysis we participated at a Tableau tutorial and one of our team members was impressed by the map visualization we got shown there. This lead to our sketch 1 from milestone 2.
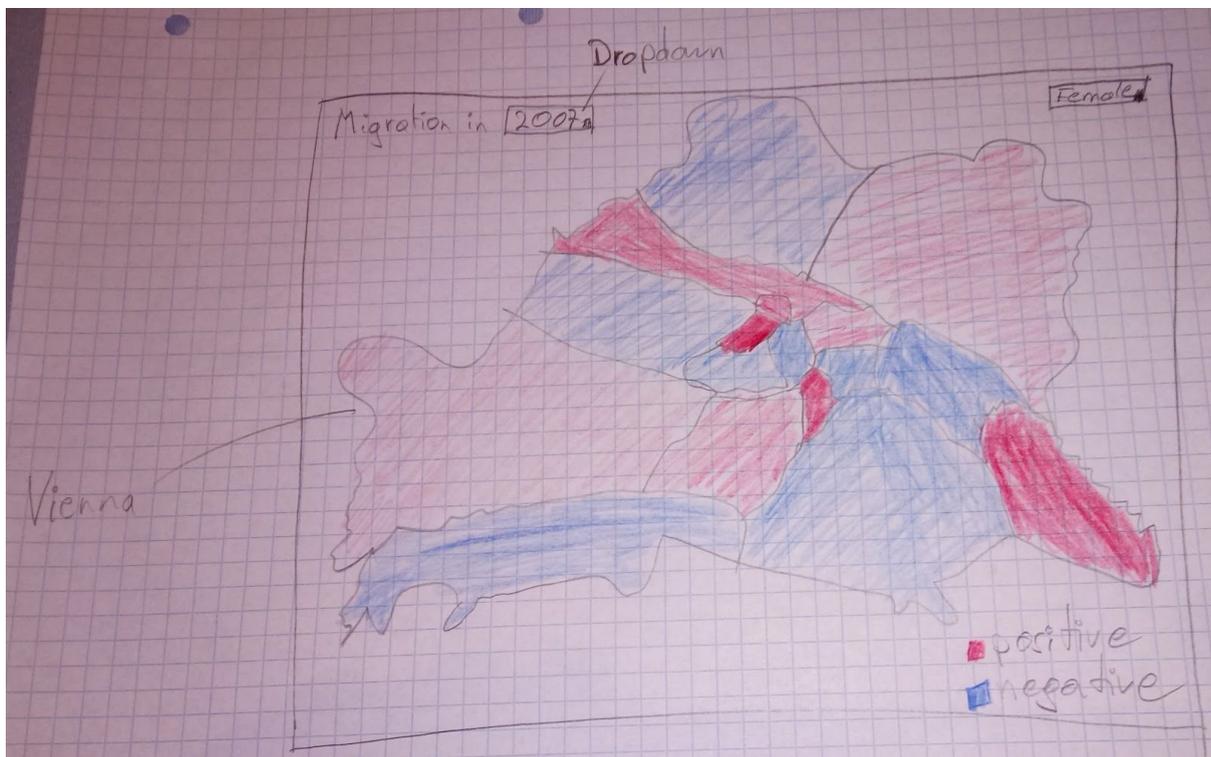

Fig. 1

This graph made it to our final version of the project. It is a heat map of Vienna's districts and their migration rates.

Besides single graphs we presented possible dashboards and a scenario of use.

3.3 Hi-Fi Prototyping
Milestone 3 was the toughest one until then. After revising we picked the best graphs from our Lo-Fi prototypes and realised them, as good as we could, with Tableau. At this point we realised how complicated it was to create some, putative easy, visualizations. For example the map of Vienna, divided in its districts, caused us quite a headache. We searched the web for possibilities and wasted some hours with background-maps but nothing was suitable for us until we got the solution in the course forum. A GEOJSON file, nearly perfect for us. Importing it into Tableau was not a big deal.

Another problem, we were criticized for in the feedback from the course instructor, was the missing labeling for each district. The reason for the missing of it are multiple. At first the datasets did not both had names, but other codes. We modified the datasets several times for different reasons, but the labelling problem was irresolvable for us. Neither District numbers from 1-23 nor district names could fit onto the map. We did not find a way to center the labeling to the center of the district, so the marking would overlap, jut out to another district or be incomplete. None of the compromises was acceptable so in the end we had to go without it at all. We faced similar problems with other charts. Lettering was always a relatively complicated thing to do.
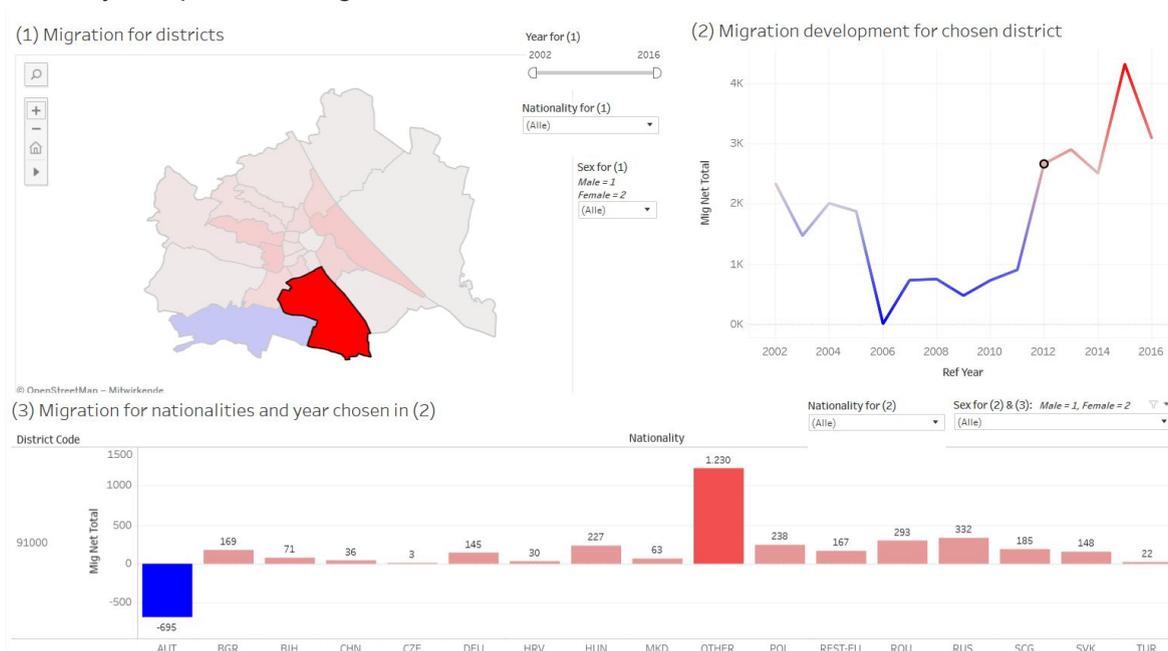

Fig. 2
This is how our final Hi-Fi prototype looked for milestone 3. The 10th district was chosen on the top left, which affected the other two graphs. On the top right 2012 was clicked which affected the bottom chart.

We chose view 1 like this because we thought its map shape invites the user to click through the data, this kind of chart is more interesting than for example a bubble chart which could represent the same data. Extremas pop out on both but with this heat map you also get an overview of Vienna and see that - in this case - outer districts have lower migration rates than those around the very center (Innere Stadt). Another side effect is that one can learn, or at least get some feeling, where districts actually are.

Deni Zukovic, 01405684   Erwin Andre, 01329095

Graph 2 is very well known by everyone who went to school. Everyone has seen this before in maths-class and knows how to read such graphs. One thing we could have done better here was the colouring. We would have liked to make all positive values red and negatives blue, but we did not manage to. In Tableau we were only able to colour it relatively, so it colours it relatively to the other data in that graph, so that small positives are often blue too.

Our bar chart 3 is should also be well known, it gives solid amount of info on little space. We chose it because it immediately gives info about various nations for the chosen district in 1. For us it was important knowledge to add to the previous graphs.

3.4 Final Submission

For the last milestone we tried to dwell on the feedback from our course instructor. We invested time in further modifying of the datasets in hope to solve the labelling problem of the heat map but did not succeed again. We added legends for the red-blue colouring which stands for low and high migration rates, changed sex-labels from 1 and 2 to male and female, shortened the district codes from e.g. 92100 to 21 and re-wrote mutated vowels to the alternative syntax.
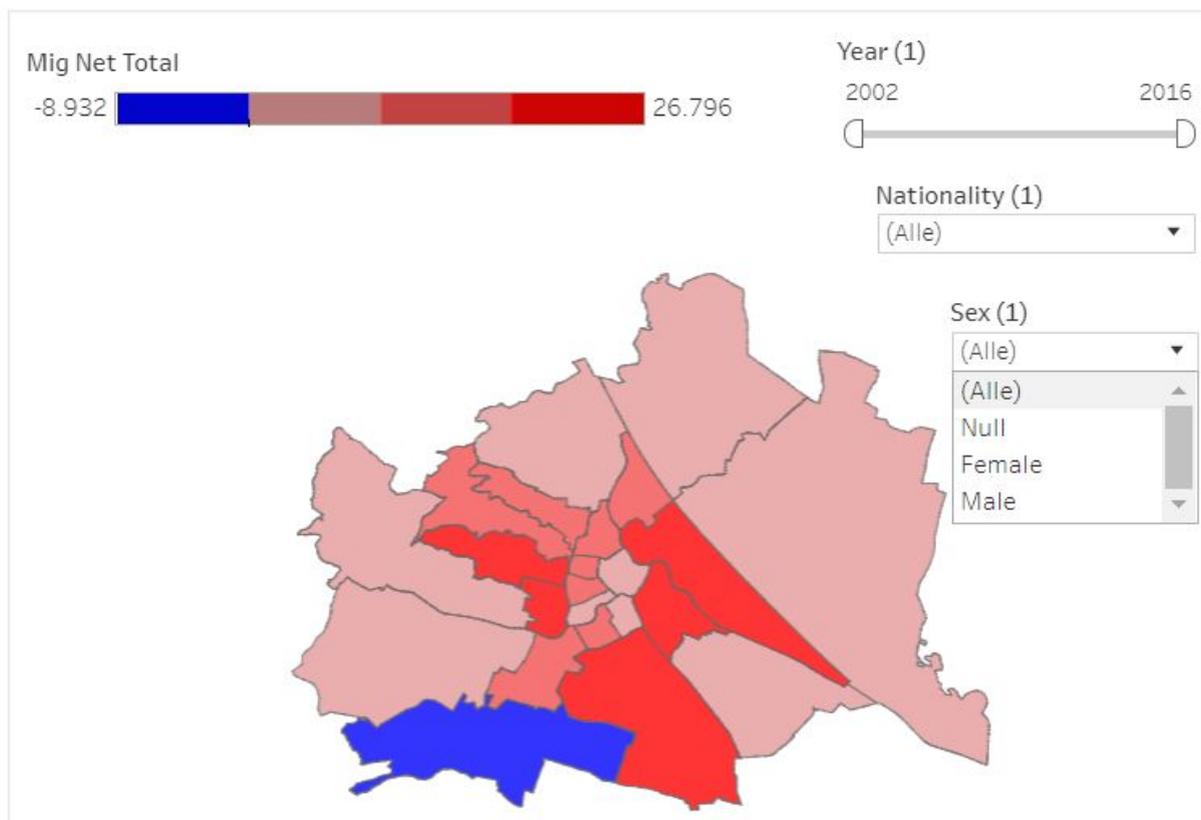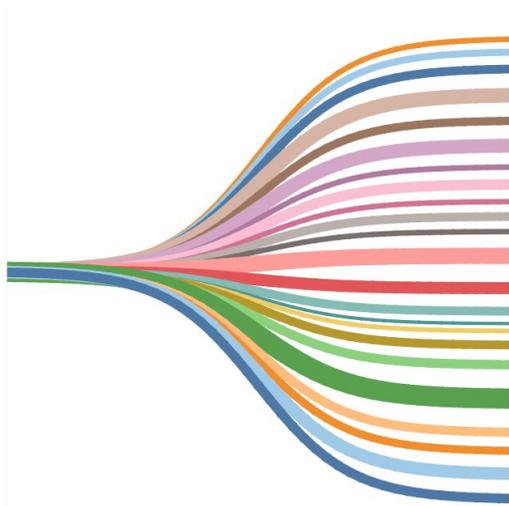

Fig. 3

Besides we created two new views, one showing generally the flow of people into which districts and the other one gives information about one specific nation moving to which districts.

Deni Zukovic, 01405684

Erwin Andre, 01329095

(4) Immigration Flow per District

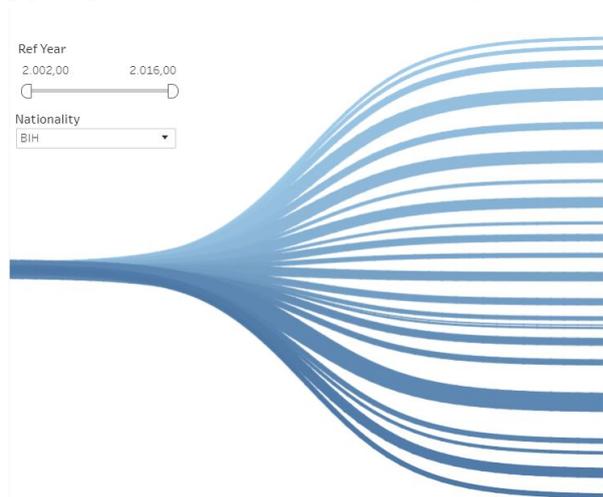(5) Immigration flow into districts for Nationality

Fig. 4

Graph 4 and 5 both give a good overview. You can see if there are extremas or as in those cases immigrants migrated pretty equally. Of course some lines are double the thickness of others but we have to consider the total population of districts. For other parameters there might be extremas that pop out more significantly.

4 Implementation

4.1 Ressources
For the implementation of our project we used Tableau. We chose Tableau, because you can get some visualization going very fast. But more complex visualizations are difficult to accomplish even in Tableau.
Our main dataset was this.
To create the map we used this GEOJSON file.

4.2 Implementation details
The map was generated with geojson data of Vienna. At this point we have to say that we were lucky finding the geojson file. Without it we would have needed to invest hours and hours of creating it ourselves or not managing doing it at all.
Another major key to success was the changing of the datasets. They needed multiple changes to work for us. The main changes we had to do were already mentioned in 3.4.

For the implementation of our sankey diagrams (which isn't featured by Tableau as mentioned before) we had to make use of the fact that you can convert data into sql. Then we had to do a lot of calculations to be be able to draw some curves. The process of understanding and transferring on how to do sankey diagrams in Tableau took some time, but we think it was worth the time.

4.3 Challenges/ Problems

We did the rest of our implementation with standard Tableau functions, sounds not that hard but figuring everything out and linking all views together was also a challenge.

The biggest challenge when working with Tableau was to import our data sources. We changed them multiple times so they were like we wanted them to fit our visualization perfectly. When we changed them Tableau didn't recognize the changes and we couldn't update them so our old views and dashboard would work with the updated data sets. We had to redo everything several times and it was very frustrating at some point.

Another challenge was to get the map right, this also took a lot of time.

5 Results

5.1 Scenario of Use

Our tool would be very helpful if you want to find out everything about for example german people where they go and in which district they have a high migration rate and also how much immigrate into which district.

So first you would change the nationality to only german on the filter near the map.
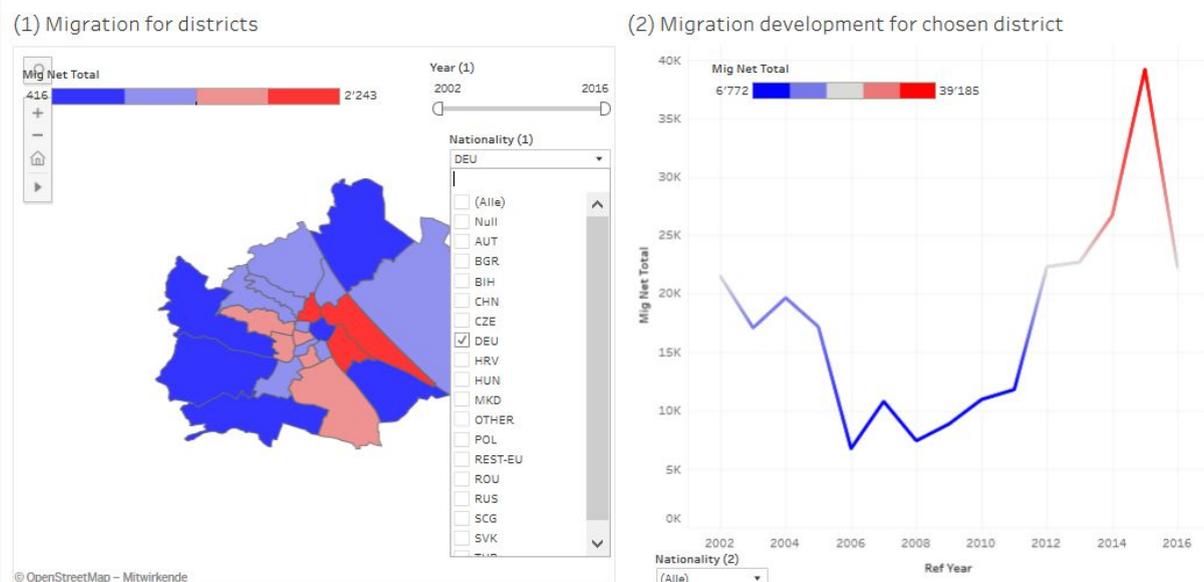

Fig. 5

Then you see a lot of germans migrate to Landstraße and Leopoldstadt, because this are the 2 districts marked red and so have the highest migration rate.

You also see on the right side that 2015 was a peak for people immigrating.

From that you can change the year and the sex if you want.

Then you would also change the nationality filter in view 5 to german to see how much immigrate into each district.
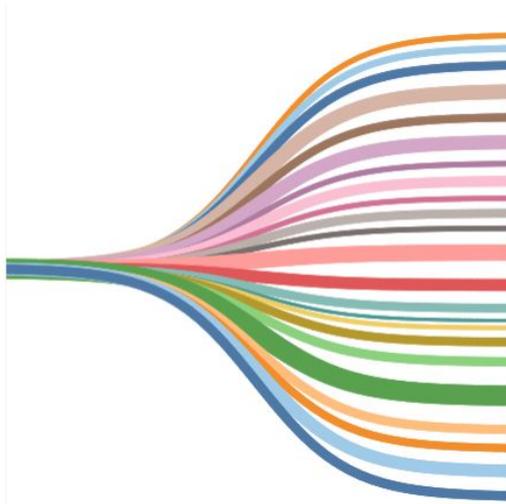
Fig. 6

You can mouse over to know what district is what curve, and bigger curves mean that there are more germans immigrating to this district than other districts.

Fig. 7

Just as in 1 we see that Leopoldstadt has a thick line which means a relatively high positive migration. Now comparing 5 and 1 might make us think that there must be an error but as we see in the legend from 1, the lowest migration rate is a positive rate of +416 and the highest is +2243, so the look of 5 makes sense. This is a problem we mentioned before, that the colouring isn't only blue for districts with more emigrants than immigrants and vice versa for red colouring. It is always in relation to the data currently showing.

5.2 Performance

When using our tool in the webbrowser you should have no performance issues, but when working with Tableau we encountered the problem that when opening the whole project or when something changes in the data it took a lot of time to load our datasets. Otherwise our project has no performance issues in Tableau too.

5.3 Feedback

We tried to implement the feedback we were given from the VIS course instructor and here are some things we changed which were requested in the feedback section:

We were able to fix all labeling issues besides the map we had before in M3, we went through our data sets and changed data if needed. We changed and fixed a lot of stuff that was problematic, in section 3.4 Final submission, there is an detailed explanation of what we were able to do.

We were able to show district names on the map but clicking on the districts wasn't possible when showing the names of each district, this must be a bug in Tableau.

Fixing our filters, so you only have one filter for selecting years we didn't realise for the final submission, because linking them all together was an issue because of different data sets. Linking datasets doesn't always work well in Tableau.

In order to make our tool better we also did some interviews with some other students, friends or family members. Here are the results we got mainly for our milestone 3 and some less for the project after milestone 4 changes.

Interview results before changes:

- + Interaction, different parameters, range of data, several views on data, easiness, colouring, possibility to choose more countries at once to combine for example countries of Former Yugoslavia,
- - filters look like "thrown on the dashboard", colouring of 2, sometimes colouring of 3, missing Syria and other countries from last flow of refugees (comment: we explained that those countries are represented with OTHER), duplicate filters

Interview results after changes (only new comments):

- + interesting style of 1,4 & 5, possibility to see details on which nationalities go to what district
- - complexity of 4 and 5, does not fit on screen

6 Discussion

6.1 Strength and Weaknesses

One strength of our implementation is, that you get to explore the data a lot, because there are 5 views on our dashboard. This might also be a weakness, because there is so much information at once. We wanted to use buttons first, for example only have 2 views in our dashboard and then click on a button to get to some other views which weren't shown before. In the end we decided against this solution, because the buttons didn't work well in Tableau Public. When clicking a button it reacted only at the second click and it opened a new tab in the browser instead of switching between two tabs we would need. These problems were not compatible with our wish to make our tool as smooth as possible. Another strength is the possibility to make several selections in our tool, that changes our graphs and visualization. You can select a time period in years or a single year, a district, nationalities and also the sex.

One Weakness of our visualization might be, that in view 4 and 5 we there was no possibility to display the district names above or in the graphs (Either the font size was to small or the name wasn't shown because there wasn't enough space). When you mouse over them you can see the district name in an quickinfo box.

6.2 Lessons learned

One of the most important things we learned during this project was that there is no perfect solution for an visualization problem. We used a lot of time to think about how we could improve our visualization and what other ways there are to make it the most comfortable for the user to explore migration flows in Vienna with our tool.

We saw that one dataset often is not enough and that you need different types of datasets to completely look on a topic. Furthermore datasets are not made for exactly for your problem and that you might need to modify them. We surely know a lot more about datasets know, we learned how to read them raw, without a interpreter and how to change them according to our needs.

We also learned that time management is very important, because we always worked on our project until 5 minutes before the deadline. We always finished in time but better time management is something we should definitely improve on.

Deni Zukovic, 01405684   Erwin Andre, 01329095

7 Who did what?

Views

	View 1	View 2	View 3	View 4	View 5
Andre Erwin	50%	0%	100%	70%	30%
Zukovic Deni	50%	100%	0%	30%	70%

Dashboards

	Dashboard 1	Dashboard 2	Dashboard 3
Andre Erwin	50%	50%	50%
Zukovic Deni	50%	50%	50%

Report

	Motivation	Rel. Work	Approach	Implementation	Results	Discussion
Andre Erwin	20%	50%	25%	75%	50%	80%
Zukovic Deni	80%	50%	75%	25%	50%	20%

8 References

https://www.superdatascience.com/yt-tableau-custom-charts-series/
https://onlinehelp.tableau.com/current/pro/desktop/en-us/maps_shapefiles.html
http://peoplemov.in/
http://www.global-migration.info/VID%20WP%20Visualising%20Migration%20Flow%20Data%20with%20Circular%20Plots.pdf
https://bensullins.com/how-to-link-tableau-sheets/
VIS Course lectures

8.1 Used data
Migration data
Map as GEOJSON