**Schedule:**

Next meeting: 3/10 2pm

Work groups:

1st group: Roza, Jorge

2nd group: Natalia, Stas, Rebecca, Boris

**Action Plan until monday**

1. Establish our limitations and challenges of working with our new path
2. New literature: possible new directions?
3. Explore new options within our new rq?
4. New thesis
5. Tableau work

**Documentation:**

26/9:

Sat together and looked critically at our dataset. Changed our plans in terms of research\ question. Found out that we do not actually have individual wages of workers, which makes answering our initial question quite difficult. We have decided to change our research question to still look at wages, but using the averaged numbers in conjunction with location data in our dataset. Our new RQ is now: “Did birthplace influence the opportunities a worker had in terms of wage and role within the VOC?”.

30/09: Group meeting.

Worked on our Tableau files - fixed issues with standardisation and linking data across different databases.

Potential issue: Relationship between voc\_places and voc\_places\_standardized is a mismatch in one of our versions of Tableau.

Started working on a map visualisation for nationalities of VOC employees. The map in Tableau uses modern borders - is this an issue?

What is relevant?

2/10:

**Our schedule until the 18th(next deadline):**



**Possible paths:**

* Correlation between ~~country of origin(landlocked?)~~/ coastal towns/ large cities
* Higher ranks: nationality and language,
* **NL or non-NL:** career opportunities, mean wages, roles in the voc,

Since we are unsure of the difficulty of linking the datasets, we decided to first focus on the simplest version of our research question. If this proves to be doable, we will expand it to one of the possible paths.

So the simplest version of our RQ is then: Were there significant differences in the wages and jobs of Dutch workers when compared to non-Dutch workers within the VOC?

Or “Did birthplace influence the opportunities a worker had in terms of wage and role within the VOC?”.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

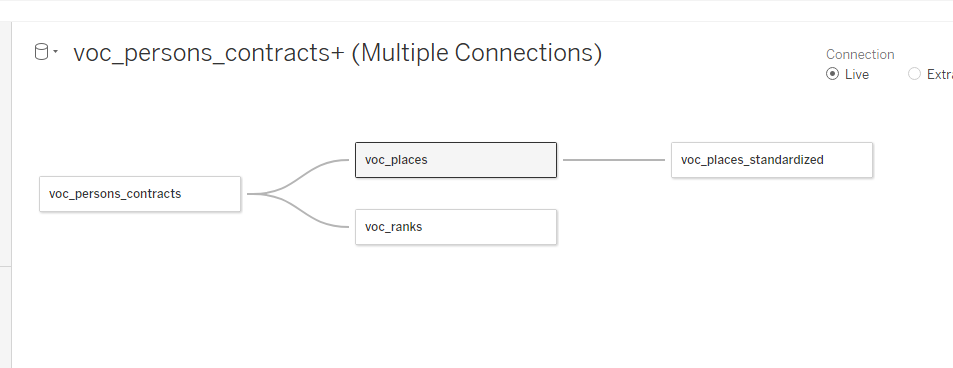
**Data sets:**

We established that the data sets are very compatible for linking, and that it was easier than expected to do so. The relational schema below closely resembles the sketch. Here is what we did:

We linked the dataset voc\_persons\_contracts to voc\_places by creating a many-to-one relationship between Place Id and Place Id. We then linked voc\_places to voc places\_standardized by creating a many-to-one relationship between the variables both named Place Standardized Id. This allowed us to see the standardized version of each person’s place of origin.

We linked the dataset voc\_persons\_contracts to voc\_places by creating a many-to-one relationship between the two variables named Rank Id. This made it possible to see each person’s median wage.

Potential problems: too many null values in e.g. wage by country graph - points to?



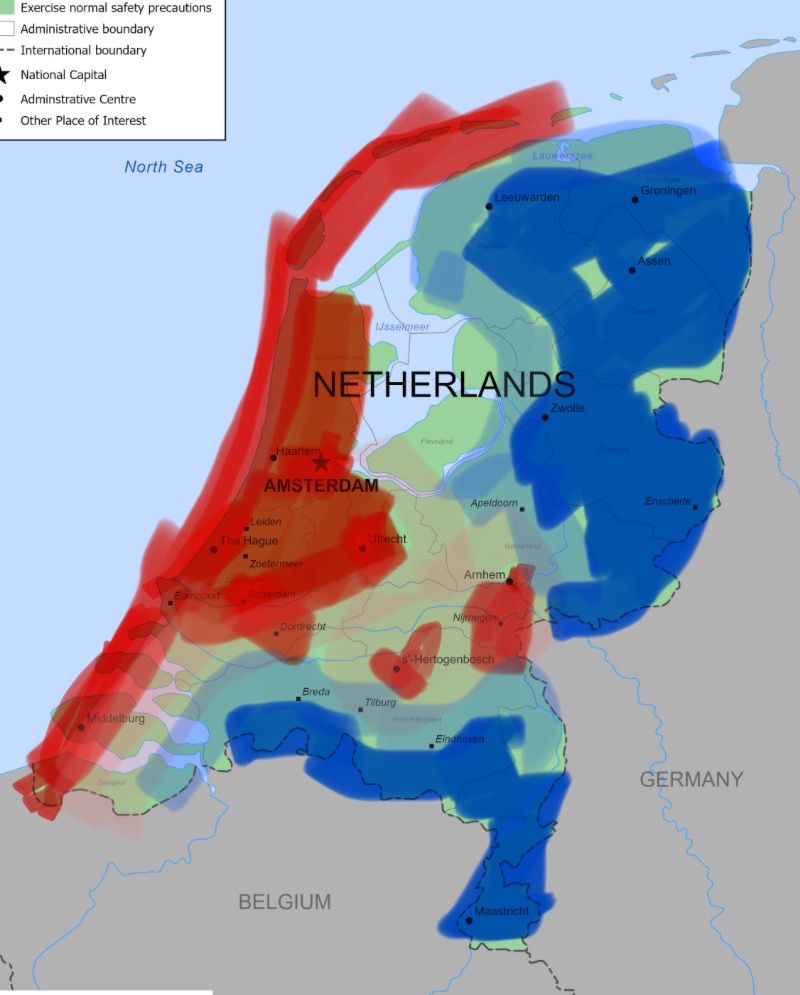
**Update 3/10:**

meeting goals:

* create subquestions to answer;
  + what visualization would be useful

what are wages impacted by?

using cities population would be troublesome - our dataset spans across 200 years and even with standardized names assigning that data might be

* heat map with time: personnel across time; 18th century only since that's when we have over 90% of data (according to the original project description)
  + use course papers on more complex visualizations for a more advanced graph
* flow chart (https://sankeymatic.com/build/) of career paths/ends?
* mean of mean wage to contrast wealthier areas of NL
* comparison with mean wages at the time outside of VOC
* hypothesis work up map for wages across NL

plan from 3/10:

* research on mean wages in NL 18th cent
* more potential visualizations and subquestions
* literature!!!!
* a research question that we’ll be happy with and can be realistically reach

Update 2.10. After consulting Erika:

We concluded we need to take another look at the cardinalities. We suppose they should all be many-to-one, instead of one-to-one as we said previously. We still need to check this. One thing we need to check: can the same person be in the dataset multiple times? Each row has a different voc id, but some have the same name. We can check whether they are from the same place but what else? In case the same person does appear multiple times, the cardinality of the relationship from voc\_persons\_contracts to voc\_ranks and to voc\_places should be many-to-many.

We should also read the paper about our datasets thoroughly, as they might explain their decisions in designing the datasets and answer our questions.

Additionally, we should analyse where our null values come from and whether they match the missing values in the datasets.

10/10:

We could look at individuals with multiple entries, whether their mean wage changed throughout time (expected yes, as experience increases).

Ambiguity because we can either look at all entries, allowing us to look at job distributions and wages. However, only looking at this obscures the data of where these individuals came from (it will count the same person each time this person occurs in the DB). So either we do use everything to answer mean wage, job distribution etc data, or we use the data that only counts every person once, which obscures mean wage and job distribution (it only considers the first entry for an individual).

Possible Literature :

1. **Robert Allen – "The Great Divergence in European Wages and Prices"**

Where:

* AIM → Analysing wage differences between Europe and Asia during the pre-industrial period and explaining the economic ‘Great Divergence’ between the West and the rest of the world.
* Wage Divergence → Between 1500 and 1750, there is a significant divergence in real wages between north-western Europe (the Netherlands and England) and the rest of Europe.
* Allen uses nominal wages and price data to construct a comparative overview between different European cities, using a cost-of-living index (‘basket of goods’).
* Quality of Life → During the 17th and 18th centuries, England and the Netherlands show an improvement in the quality of life and a greater variety of consumer goods, despite falling wages in other European cities.
* Welfare Ratios → Allen uses the concept of the ‘welfare ratio’ to compare average income with the cost of a basket of goods needed by a household, showing how purchasing power varied between different European cities.

1. **Els Jacobs' *"Merchant in Asia"***

***Labor and Wages in VOC***

* Els Jacobs' analysis reveals the VOC’s hierarchical labor system, showcasing significant wage disparities based on factors like employee rank, skill level, and geographical location of employment (Jacobs, 2000, 112). Workers stationed overseas typically earned more due to the hazardous and unhealthy conditions they faced compared to their counterparts in the Netherlands.

***Wages for Sailors and Soldiers***

* Sailors constituted a significant portion of the VOC workforce and earned modest wages. Many supplemented their income through private trade, known as "ventjagers," which allowed them to buy goods in Asia and sell them in Europe (Jacobs, 2000, 114). Soldiers, tasked with protecting VOC assets, also received low pay but often earned higher wages for serving in particularly dangerous or remote locations (Jacobs, 2000, 116).

***Administrative and Skilled Labor in Asia***

* Higher-ranking administrative positions in Asia offered more substantial salaries and benefits, including housing allowances and provisions (Jacobs, 2000, 118). These positions, particularly those held by upper-level officials, provided opportunities for significant personal wealth through private trade activities beyond their official duties (Jacobs, 2000, 120).

***Labor Exploitation and Wage Inequality***

* The VOC’s labor system was characterized by stark wage inequalities, where high-ranking officials amassed considerable wealth compared to ordinary workers struggling to meet basic needs (Jacobs, 2000, 123). The VOC also exploited local labor in Asia, paying them significantly less than European workers, which enabled the company to maintain lower operational costs while exerting control over its territories (Jacobs, 2000, 125).

1. **J.L. van Zanden (1999): "Wages and the Standard of Living in Europe, 1500–1800"**

* **Urban vs. Rural Wages**
  + The economic landscape of the Dutch Golden Age highlighted a significant wage disparity between urban and rural workers, with urban centers like Amsterdam and Rotterdam offering much higher wages due to thriving industries such as shipbuilding and international trade (Van Zanden, 2009, 185). Rural laborers often engaged in subsistence farming, earning much less and benefiting less from the VOC’s wealth.
* **Skilled vs. Unskilled Labor**
  + There was a notable gap in wages between skilled and unskilled labor, driven by the VOC’s demand for specialized skills in shipbuilding and maintenance. Skilled workers, such as artisans and craftsmen, commanded significantly higher wages than their unskilled counterparts, whose wages remained stagnant (Van Zanden, 2009, 187).
* **Gender Wage Gap**
  + The gender wage gap was pronounced during this period, as women predominantly occupied lower-paying jobs such as domestic work and textile production. Despite performing similar tasks as men, they were often paid less, reflecting societal norms that undervalued female labor. The male dominance in the VOC workforce further entrenched these disparities (Van Zanden, 2009, 191).
* **Impact of VOC on Wage Inequality**
  + The VOC significantly influenced wage structures in the Dutch Republic, leading to economic growth but also exacerbating wage inequality. High-ranking officials in the VOC often amassed considerable personal wealth through lucrative salaries and opportunities for private trade (Van Zanden, 2009, 198). In contrast, ordinary sailors and soldiers faced harsh conditions and low wages, limiting their economic mobility (Van Zanden, 2009, 200). The economic benefits of VOC trade primarily flowed to wealthy merchants and investors, deepening the divide between the urban elite and working class (Van Zanden, 2009, 203).
* **Long-Term Trends in Wage Inequality**
  + Van Zanden contextualizes wage inequality in the Netherlands within broader European economic trends. The VOC’s trade and capital accumulation were key drivers of the Great Divergence, where Western Europe began to economically outpace other regions (Van Zanden, 2009, 210). This concentration of wealth in the hands of a few elite families, coupled with stagnant wages for the working class, laid the groundwork for future social unrest and economic challenges in the transition to the modern industrial era (Van Zanden, 2009, 215).

10/10:

**RQ: “To what extent did the geographical origin of VOC workers influence their wages and career opportunities within the company?"**

**Questions about wage distribution inside the Netherlands (internal):**

1. How did wages actually work within the VOC? And how important was the company for the Dutch economy? (Literature question).

The VOC was massively important for the Low Countries both economically and for

its impact on Dutch society. For instance, it swept the streets clean of homeless and

unemployed people as it needed so many workers in lower paying roles. It created a

huge surplus of women in Dutch society as close to a million (mostly) men left for the

Indies of which only about half returned. Despite only having 6% of the Dutch

merchant ships, the VOC accounted for a quarter of Dutch sailors and a quarter of the

miles covered by Dutch ships (Poeze, 1996, p. 381).

“These Europeans were the in-loco managers of these far flung trading empires. Their pay accounted for the jobs’ distant location—often affected by military conflict and ravaged by disease—, which would only be reached if they survived the original voyage. Nevertheless, these were amply sought jobs given their long-term contracts and the prevalence of poverty and unemployment in pre-industrial Europe. Overseas jobs were also opportunities to make a fortune and return to Europe with an improved social and financial standing, for those lucky enough to survive.” (Rei, 2013, p. 28). <- This excerpt explains *why* European workers chose to do this job, despite the bad circumstances and prospects.

1. Did the workforce benefit when the VOC was doing well? Did their wages rise as the VOC became more powerful? (Literature question).

Labour compensation within the VOC was structured so that employees typically had a fixed wage and an above wage income, which typically came in the form of trade privileges, gifts and pays from local jurisdictions or financial advantages associated with a given position (Rei, 2013, p. 41). Rei (2013, p. 41) also proposes that just 64% of total compensation was spent on wages, leaving the remainder to these differently structured wages. We can imagine how employees from the VOC would actually benefit from the VOC doing well as a company. Especially those in higher-paying positions.

1. How do correlations between average wages and biggest cities emerge?

**Questions about wage distribution of the NL compared to non-NL (external):**

1. Comparison on ranks’s categories between Dutch and non-Dutch workers within the VOC

Also the Pusztai & Teszelszky paper that is discussed below.

1. Dutch non dutch wage comparison - no need for visuals

Pusztai & Teszelszky (2016, p.28) speak about how foreign workers were necessary to make up for the massive amounts of people that it had to employ to keep running - 973.000 people left for the Cape of Good Hope and only one third ended up returning. Foreign workers were recruited to make up for the huge numbers, often being employed in the lower ranks: Of all soldiers about 60% were foreign, for sailors this number was close to 40%. The authors say that many of the seafarers were fired, killed or went missing. Some others were sentenced to die for crimes committed out of boredom while in Asia and waiting for their ships to return.

An interesting little piece from the article (p.39) to sketch the circumstances:

“The tropical adventure ended sadly for most VOC servants. Instead of wealth and happiness, they usually found death in the East. The death rates were particularly high in the period after 1700. That was also the period when the VOC, due to a lack of its own Dutch labour force, attracted more and more people from abroad. According to Gaastra around 1770, 80% of all soldiers and 50% of all seamen employed by the Company came from abroad. Especially the lower personnel, i.e. soldiers and sailors, had little chance of survival. Of the soldiers who in 1775 had survived the voyage to Batavia, 80% died within two years. These high mortality rates were mostly due to malaria.”

1. Rank advancement over time (using the repeating ppl)
2. How did the wages differ between workers who were working in NL and the ones who worked overseas?

Workers based in Asian ports, for instance, often received higher wages compared to their counterparts in the Netherlands due to the dangerous and unhealthy conditions associated with overseas work (Jacobs, 2000, 112).

**Dutch workers** employed in the Netherlands (e.g., craftsmen, shipbuilders) generally earned less than those working overseas, despite having specialized skills. The jobs based in the Netherlands, while stable, did not offer as many opportunities for high earnings.(Jacobs, 2000, p. 112).

**Non-Dutch workers**, particularly those stationed in Asia, often received higher wages due to the dangerous and unhealthy conditions abroad. Sailors and soldiers, both Dutch and non-Dutch, were paid premiums for hazardous postings such as in Batavia or Ceylon. However, local Asian workers employed in VOC territories were paid significantly less than Europeans.(Jacobs, 2000, p. 125).

1. How did the VOC exploit local labor in its Asian operations, and how did the treatment of local workers differ from that of European employees?

The VOC heavily relied on local, often bonded or enslaved, labor in its Asian operations. These non-Dutch laborers were paid little or nothing, in stark contrast to their European counterparts, contributing to significant wage inequality within the VOC's workforce (Jacobs, 2000, p. 125).

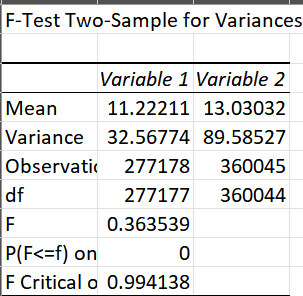
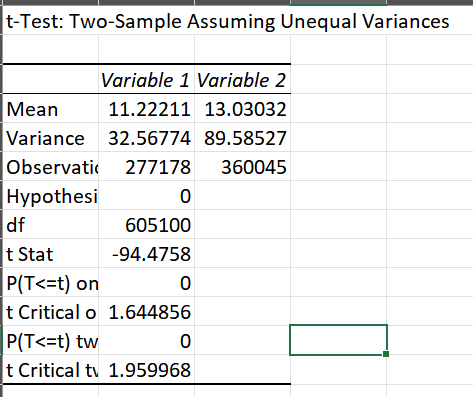
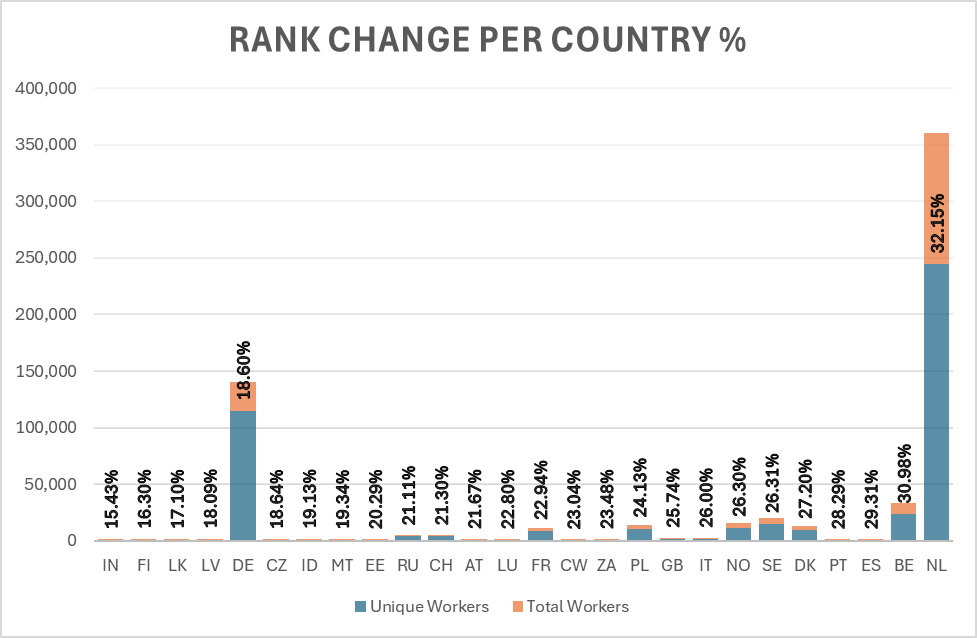
**Visualizations:**

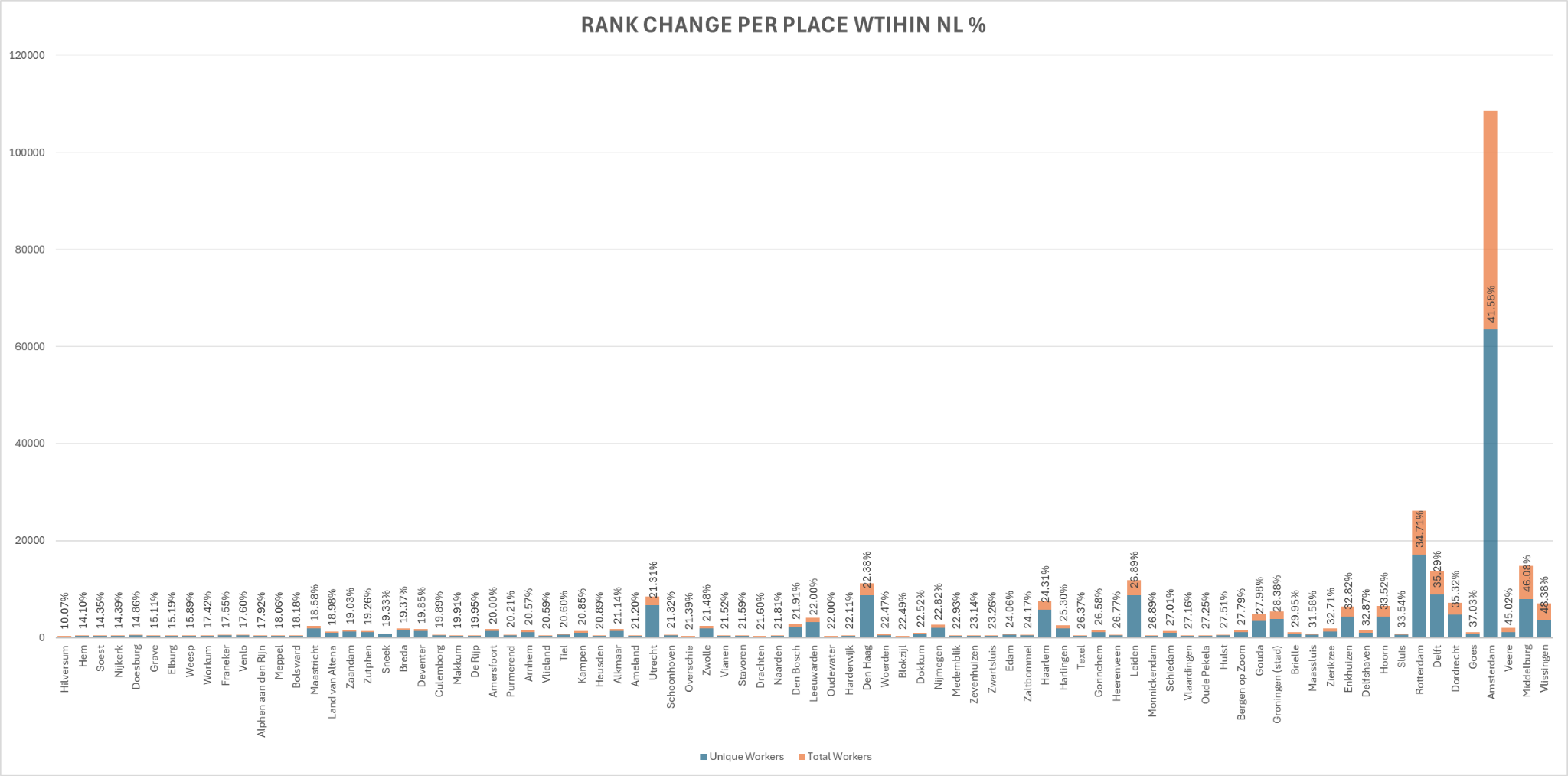
* **Heat map** of avg wages in cities of nl
  + Global map zooming into NL???
* **2 pie charts**: Rank category differences
* Numerical statistics and significance data

Workflow Steps:

* Do we exclude Indonesian/Asian employees since the majority of that workforce is excluded since they were enslaved and only the highest ranks are included in our data?
* Should we use the 30 ppl minimum when comparing international stats
* differences between Dutch and Non-Dutch VOC employees could be a subquestion instead of being included in the RQ
* We will choose to use a visualisation of a heatmap of average wage for different geolocations in our poster/presentations. There will be a zoomed out picture depicting a slice of the world (mostly Europe) to grasp the visualisation and then dive into the detailed part of Netherlands area showing our interesting data.
* We found a paper that can support our claim about the difference between urban workers and rural workers.
* Find if statistical difference between mean wages inside and outside of the netherlands with R.
* Find how many ppl changed roles per place by comparing unique vs normal person.
* Do workers from the more farming-focused areas differ from the rest?
* What were the VOC rules/policies on this topic? Find out from papers
* inland populations working in farming or
* better jobs for dutch - bc of social capital?

**Statistical Significance of Findings Analysis:**

* First I made a wage-vocop Id database out of the linked databases on Tableau.
* Filtered it by country code and excluded NL and did the opposite later (excluding null values in both cases).
* Exported both of these databases into Excel where I performed an F-test two sample for variance equality at 5% significance level. With these results:
* P value is effectively 0 which is not close to the significance value established by our Alpha. This means that, when comparing our dataset’s means, we will use Welch’s t-test.
* The two sample t-test assuming unequal variances at a 5% significance level yields these results:
* We can see the mean of the 2 variables are different but are they significantly different? The t-value of roughly -94.5 is considerably below the two tail critical value of roughly 1.96. This signifies grat statistical differences between the 2 means.
* From this we can conclude dutch were privileged compared to the rest of the world combined when it came to ranks within the VOC.
* We noticed there were a lot of people in the database that were counted multiple times since they changed ranks so we made a new calculated field in Tableau using this formula: { FIXED [Person Cluster Id] : MIN([Vocop Id]) }. This gave us a new parameter we called Unique Persons.
* By Comparing unique people to contracts overall, we can roughly see rank progression over time given that we assume that most people are not getting demoted or switching for worse.
* We then categorised this data by country without linked datasets on Tableau and exported it to excel. There, we computed the percentage change in ranks per country and plotted a stacked bar chart demonstrating the number of workers vs unique workers tagged and sorted by the % change. We excluded countries with less than 250 workers since it led to outliers.
* In this chart, we can see that even though there seems to be a correlation between how many workers a country has in the VOC and its proportionate rank changes, there are outliers like Germany and perhaps more importantly, it is a great privilege to be a Dutch worker as it gives you almost ⅓ chance to move up the ranks in your stay at the VOC.
* This prompted us to dive deeper within the Netherlands and perform a similar analysis inside of the modern boundaries of the country.
* For this one we also filtered by at least 250 workers per place and computed the results in the same type of graph (clearer in presentation):



* These results show that higher populated areas like Amsterdam and Rotterdam provide a lot of opportunities for workers to move up in ranks and make a name for themselves. However, more importantly it shows us there is an area in the middle peninsula of Zeeland that has a disproportionately high amount of rank promotion.

Let's get this going before the next meeting on the 15th(?):

Presentation content

* An Introduction with a brief background about the project (e.g., context); STAS
* A short explanation of the project’s objectives (i.e., RQ, thesis statement); STAS
* The project's workflow;
* A description of the dataset including limitations; (Boris)
* A description of the methodology, including Work Packages, aka division of tasks;
* A discussion of your findings (analysis and interpretation of the results);
* A discussion of the relevance of your findings, limitations and future work;
* A brief description of the team.

Documentation to do:

* Project Overview
* Data Acquisition
* Methodology
* Workflow Steps
* Challenges and Solutions
* Ethical Considerations (excluding Asia - give our reasoning)
* Results
* Documentation and Sustainability
* Reflection

PLAN FOR NOW:

Nat: workflow of our entire project (waiting for the notes on our decisions in visualizations and stat analysis)

Staś: Project overview and presentation.

Boris: Dutch internal VOC literature review, Dutch and English studies; dataset limitations

Roza: Does the rank improve over time(by country?)

References for literature review

Pusztai, G. & Teszelszky, K. (2016). In Dienst van de VOC. Een voorlopige inventarisatie van Hongaren in dienst van de Verenigde Oost-Indische Compagnie (1602 - 1795). *Acta Neerlandica, 12,* 25-108.

Poeze, H. A. (1989). Korte signaleringen. *Bijdragen Tot De Taal- Land- En Volkenkunde /*

*Journal of the Humanities and Social Sciences of Southeast Asia*, *145*(2), 379–392. <https://doi.org/10.1163/22134379-90003263>

Rei, C. (2013a). Careers and wages in the Dutch East India Company. *Cliometrica*, *8*(1),

27–48. https://doi.org/10.1007/s11698-013-0093-3

### The presentation should contain:

### The presentation should contain: