Project Proposal Group One:

Lisanne Koek, Samantha Mooij, Sierra Scott

**Research question**: Is there a correlation between an individual’s initial rank category (specifically, ‘Military’, ‘Sea’, or ‘Trade’) within the Dutch East India Company (VOC) and their mortality risk, and if there is, what does this indicate about the working conditions and systemic inequalities within the VOC during the eighteenth century?

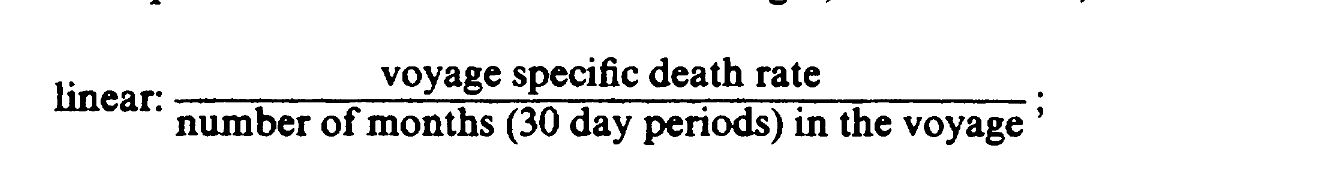
**Thesis statement:** Mortality rates by initial rank category (Military, Sea, and Trade) in the eighteenth-century Dutch East India Company (VOC) may provide insight into how rank-based working conditions increased the vulnerability of lower-status employees and reveal patterns of systemic inequalities and coercive labor practices that structured the VOC.

**Literature**:  
J. de Hullu (1913) documents high disease and mortality on VOC ships, noting Military personnel were often affected first and worst, supporting the idea that mortality varied by rank. Similarly, Worden (2009) examines the 1732 Loenderveen voyage, with a 35% mortality rate instead of the 8% average. Scurvy disproportionately affected lower-ranking sailors and soldiers, while officers were somewhat protected. Voyage timing, pre-existing malnutrition, and tensions aboard increased risks, and harsh disciplinary practices added to fear among inexperienced crew. Van Rossum (2019) places the VOC within the context of early modern globalization, highlighting mobilization of diverse labor across ports and long-distance transport. Workers from different social, regional, and cultural backgrounds were brought together under complex hierarchies and subjected to legal and disciplinary regimes, illustrating how systemic inequalities and coercive labor practices structured work in the company. Together, these sources provide context for interpreting mortality patterns found in the dataset (de Hullu, 1913; Van Rossum, 2019; Worden, 2009).

**The dataset**  
The dataset covers European VOC personnel who signed contracts in the Netherlands between 1600 and 1794. It records names, places of origin, career trajectories, voyages, days in service, and reasons for not renewing contracts. Derived from VOC pay ledgers, it was digitized by historian Ton van Velzen (2000–2012). Because the ledgers were organized by individual contracts, careers were initially difficult to trace. From 2017 to 2024, a collaborative project of historians and data scientists restructured the data, disambiguating individuals, standardizing place names, and categorizing ranks. The enriched dataset contains 460,452 disambiguated persons across 774,200 contracts (Petram et al., 2024).

**Data limitations**A major limitation is the dataset’s “Dutch bias.” VOC clerks were more familiar with the Dutch language, which reduced variation of names and places of origin of Dutch personnel. During standardization (Petram et al., 2024), resources focused on the Netherlands, reinforcing this bias. Places of origin were often assigned through assumptions that favored larger settlements, skewing representation. Contract linking was deliberately conservative, reducing false positives but increasing false negatives, making careers seem shorter and deaths more frequent. This means that people with common names may appear as having only one voyage ending in death. While this may not distort rank-based comparisons, overall mortality is likely overstated. Reliability also varies by rank. The data for lower ranks, such as sailor and soldier, are more common and therefore their outcomes are more reliable. Officers were rarer, meaning that each missing case carries a disproportionate weight.

**Processing our dataset**We will use the R coding language due to group familiarity with the language. To inspect mortality, we will work with the ‘reason\_end\_contract’ variable, as this indicates whether a worker’s contract was suspended because they were deceased. This categorial variable has no missingness and contains 32 categories representing different termination reasons. To thoroughly answer our research question, we anticipate converting this variable into a binary variable that indicates whether or not an individual had their contract terminated due to being deceased. To inspect the different ranks within the VOC, both the “voc\_persons\_contract” and “voc\_ranks” sets will be used. Since ‘rank’ is a categorical variable with several levels in both datasets, we plan to group the ranks into larger categories such as “shipmen” or “military” to minimize the number of levels and more thoroughly address our research question. As of now, we think it would be optimal to transform the ‘rank’ variable into a categorical variable with a maximum of 5 levels. Mortality rates should be standardized according to length of voyage according to Riley (1981). All restructuring and standardization will be done within the dataset, so no external linking is needed.



***Group Charter***

***Goals***

* Sierra: as a U.S. citizen, would like to learn more about the historical impact of the Dutch VOC and uncover valuable insights about VOC workers.
* Lisanne: would like to learn how to work with quantitative data and experience working with people from different academic backgrounds.
* Samantha: would like to learn how to use historical data and what insights it can provide.

***Backgrounds and Strengths***

* Sierra: has data science background and is familiar with predictive modelling and utilizing data visualization tools in the R coding language
  + Can create valuable data visualizations such as bar charts, scatterplots, etc.
* Lisanne: studies Media, Art, Design, and Architecture (with a specialization in art) and is familiar with working with historical documents.
  + Is good at literature reviews and essay writing.
* Samantha: sociology, good at making connections, writing and literature research. Also has a background in data analytics.

***Pressures***

* Different schedules and work speeds could cause imbalance, making clear communication and planning essential.

***Communication strategies***

* Outside the classroom we communicate via WhatsApp and online meetings.

***Meetings***

* We will meet weekly, with additional sessions added if needed.

***Planning***

* Everyone familiarizes themselves with the dataset by September 16
* Online meeting on September 17 to discuss ideas and divide tasks

Bibliography

De Hullu, J. (1913). ZIEKTEN EN DOKTERS OP DE SCHEPEN DER OOST-INDISCHE COMPAGNIE. *Bijdragen Tot de Taal-, Land- En Volkenkunde van Nederlandsch-Indië*, *67*(2), 245–272. <http://www.jstor.org/stable/20769694>

Petram, L., van Velzen, T., van Rossum, M., & Huygens Institute. (2024). *VOC Opvarenden: Digitale editie van de scheepssoldijboeken, 1600–1795*. Amsterdam: Huygens Institute.

Riley, J. C. (1981). *Mortality on long-distance voyages: The experience of the East India Company, 1600–1800*. Ann Arbor: University of Michigan Press.

van Rossum, M. (2019). Labouring Transformations of Amphibious Monsters: Exploring Early Modern Globalization, Diversity, and Shifting Clusters of Labour Relations in the Context of the Dutch East India Company (1600–1800). *International Review of Social History*, *64*, 19–42. <https://www.jstor.org/stable/26646386>

Worden, N. (2009). ‘Below the Line the Devil Reigns’: Death and Dissent aboard a VOC Vessel. *South African Historical Journal*, *61*(4), 702–730. <https://doi.org/10.1080/02582470903500384>