Interactive Visual Exploration of Longitudinal Historical Career Mobility Data

HKUST VisLab

Yifang Wang[1,2]   Hongye Liang[2,3]   Xinhuan Shu[1]   Jiachen Wang[2,3]   Ke Xu[4]

Zikun Deng[2,3]   Cameron Campbell[1]   Bijia Chen[5]   Yingcai Wu[2,3]   Huamin Qu[1]

1 香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

浙江大學 ZHEJIANG UNIVERSITY

3 之江实验室 ZHEJIANG LAB

4 NYU

5 中國人民大學 RENMIN UNIVERSITY OF CHINA

Project Background

Quantitative Historical Datasets
CGED-Q
Career Mobility

- Digital Humanities
- History
- Sociology
- Demography
- ...

- Career
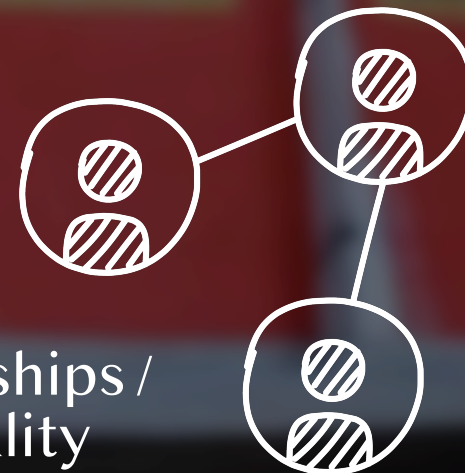- Health
- Family
- Migration
- ...

Historical Quantitative Datasets

China Government Employee Database-Qing (CGED-Q)

- One-year project
- CGED-Q records the career trajectories of over 340,000 government officials in the bureaucracy of Qing China from 1760 to 1912

- Career Mobility: the study of career trajectories and the factors influencing them.

Social Ladders

Social Relationships / Social Inequality
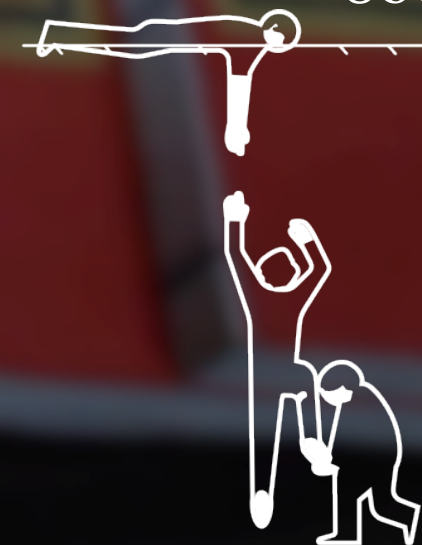
Existing Career Mobility Analysis

Hout, 1983

Father's Occupation	Son's Occupation					
	Upper Nonmanual	Lower Nonmanual	Upper Manual	Lower Manual	Farm	Total
Inflow Percentages						
Upper Nonmanual	34.5	17.6	11.5	8.1	1.8	14.7
Lower Nonmanual	17.7	17.7	9.7	8.8	2.1	11.3
Upper Manual	19.5	21.9	32.6	21.1	4.8	20.5
Lower Manual	18.4	30.8	29.4	41.8	10.5	30.1
Farm	10.0	12.0	16.8	20.2	80.9	23.4
Total	100.0	100.0	100.0	100.0	100.0	100.0

Jarvis and Song, 2017

Table 4. Log-Linear Model Fit Statistics

Model	df	L^2	p-value	BIC	Δ
Men (N = 56,200)					
Baseline Models					
0a: Origin + Destination + Period	27,972	244,675	<.01	−61,245	69.0
1a: 0a + Origin×Period + Destination×Period	27,380	239,997	<.01	−59,449	68.5
2a: 1a + Parameter-ABCD	27,292	27,431	.28	−271,052	18.2
Period-Varying Topological Mobility Models					
3a: 2a + Period×Parameter-A	26,992	24,874	>.99	−270,328	14.2
4a: 2a + Period×Parameter-AB	26,956	24,723	>.99	−270,086	14.1
5a: 2a + Period×Parameter-ABC	26,944	24,661	>.99	−270,017	14.1
6a: 2a + Period×Parameter-ABCD	26,940	24,648	>.99	−269,986	14.1
7a: 2a + Period×Parameter-A'B'C'D'	27,276	25,659	>.99	−272,650	16.0
Women (N = 47,180)					
Baseline Models					
0b: Origin + Destination + Period	27,972	174,593	<.01	−126,434	66.6
1b: 0b + Origin×Period + Destination×Period	27,380	166,191	<.01	−128,465	65.7
2b: 1b + Parameter-ABCD	27,292	19,635	>.99	−274,074	15.2
Period-Varying Topological Mobility Models					
3b: 2b + Period×Parameter-A	26,992	18,287	>.99	−272,193	12.3
4b: 2b + Period×Parameter-AB	26,956	18,154	>.99	−271,939	12.2
5b: 2b + Period×Parameter-ABC	26,944	18,102	>.99	−271,862	12.1
6b: 2b + Period×Parameter-ABCD	26,940	18,077	>.99	−271,844	12.1

- Social scientists
  Group-level analysis

Existing Career Mobility Analysis

Hout, 1983

Jarvis and Song, 2017

- **Hypothesis-driven and ill-suited for exploration**

- Social scientists
  Group-level analysis

Existing Career Mobility Analysis

Hout, 1983

Jarvis and Song, 2017

- **Hypothesis-driven and ill-suited for exploration**

Qu et al., 2016

Fig. 8. The multi-task and multi-feature prediction model

Xu et al., 2015

- Social scientists
  Group-level analysis

- Data scientists
  Individual-level analysis

8

Existing Career Mobility Analysis

Jarvis and Song, 2017

- Hypothesis-driven and ill-suited for exploration

Qu et al., 2016

- Do not target career mobility -- emphasizes the characteristics of groups and more complicated social relationships

- Social scientists
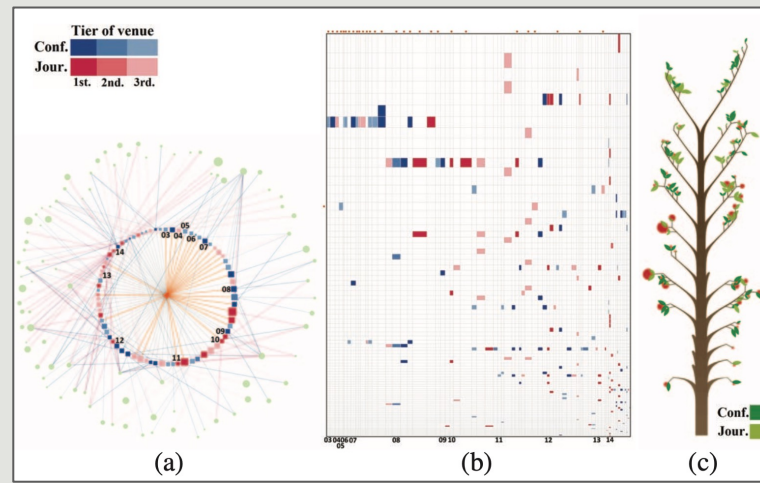 Group-level analysis
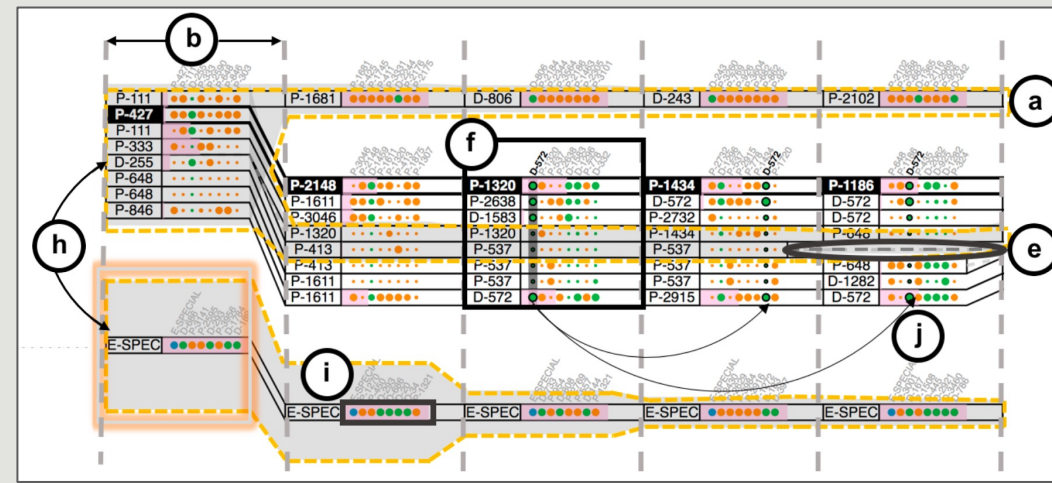
- Data scientists
  Individual-level analysis

Existing Career Data Visualization (1/2)

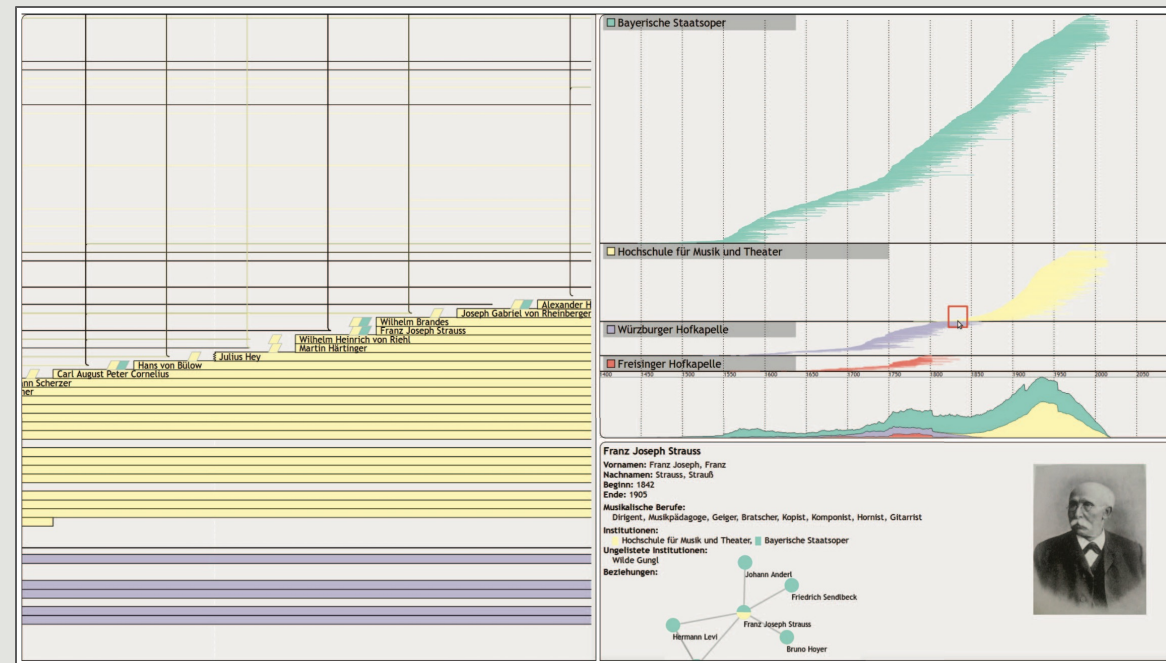Fung et al., 2016

Guo et al., 2018

Du et al., 2016

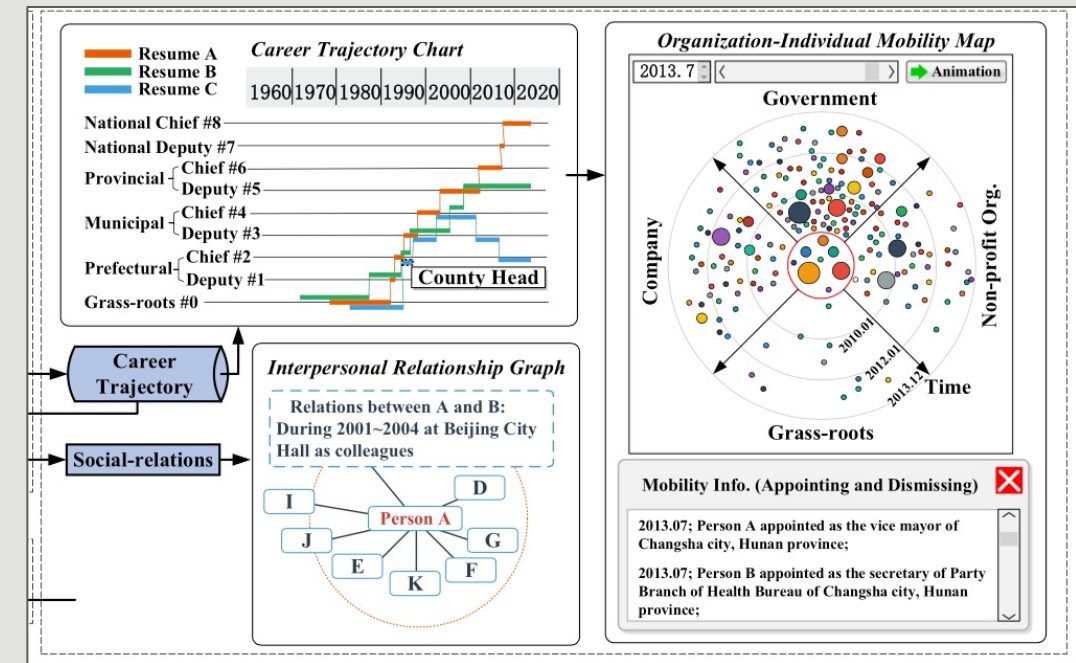- Network summarization
- Sequence summarization
- Similarity comparison

Existing Career Data Visualization (2/2)

Khulusi et al., 2019

Zhang and Wang, 2019

- Multi-task analysis

Existing Career Data Visualization (2/2)

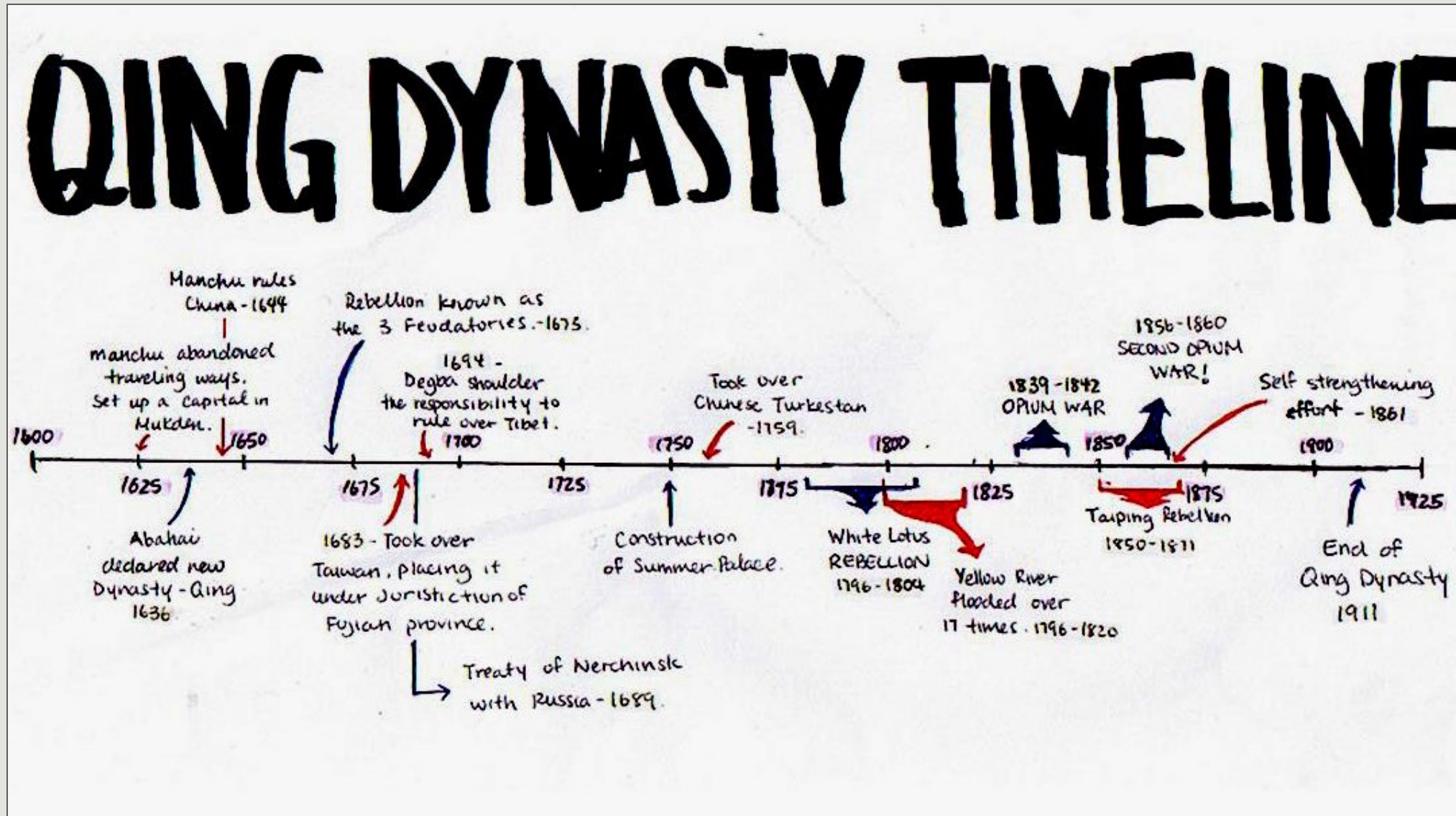Khulusi et al., 2019

Zhang and Wang, 2019

- **Support a limited number of career comparisons and lose the overall context**
- **Are limited by the short time range**

- Multi-task analysis

Data Description

Data Description

- **Timestamp**: The year and season covered by the record
- **Name**: The official's real name in the Qing dynasty
- **Unique ID**: A 12-character unique identifier of each official generated by experts

Personal Info

Job Info

Data Description

- **Timestamp**: The year and season covered by the record
- **Name**: The official's real name in the Qing dynasty
- **Unique ID**: A 12-character unique identifier of each official generated by experts
- **Birthplace**: The geographic origin of the official
- **Family Background**: A identity indicating whether the official was associated with the imperial lineage
- **Ethnicity**: Three types of officials are identified based on ethnicities: Manchu, Mongol, or Han
- **Exam Degree (科举结果)**: The examination or purchased degree held by the official. Those with high examination degrees were political elites

- **Job Location**: The geographical location of the official's current job
- **Job Department**: The department in the bureaucracy where the official works. We classified them into fifteen categories according to experts' suggestions
- **Job Level (品级)**: The administrative rank of the job in the bureaucratic hierarchy, represented by a number (ranging from 10 to 1 with 0.5 as a step).

16

Challenges (1/3)

- How to visualize a large volume of longitudinal career data with a complex data structure?
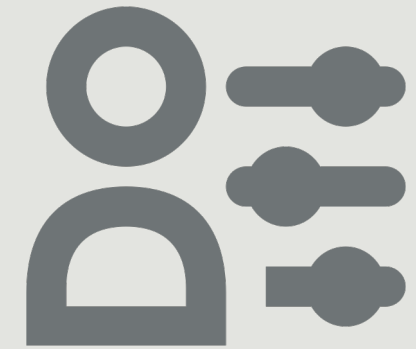
- Temporal
  - 1760 ~ 1912

- Network
  - Colleagues
  - Townsmen
  - Classmates

- Multi-attributes
  - Personal Info
  - Job Info

Challenges (2/3)

- How to visualize a large volume of longitudinal career data with a complex data structure?
- How to extracte and highlight social groups and social relationships from this large dataset?

- Predefined Group vs. Latent Group

- Semi-automatic

Challenges (3/3)

- How to visualize a large volume of longitudinal career data with a complex data structure?
- How to extracte and highlight social groups and social relationships from this large dataset?
- **How to support multi-level mobility analysis and reasoning?**

- Overall Level

- Group Level

- Individual Level

Tasks and System Design

Tasks
Glyph
Population Flow

Experts and Tasks

- Experts

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY | SCHOOL OF HUMANITIES AND SOCIAL SCIENCE

中國人民大學 清史研究所 THE INSTITUTE OF QING HISTORY

Overall

- What are the general characteristics of career mobility?
- What special features do the groups with vertical movements have at different time periods?

Group

- What are the characteristics of different social groups?
- What is the mobility pattern for each group?

Individual

- What are the mobility characteristics for different individuals?
- How do the mobility patterns of each individual and his social relationships change over time?

CareerLens
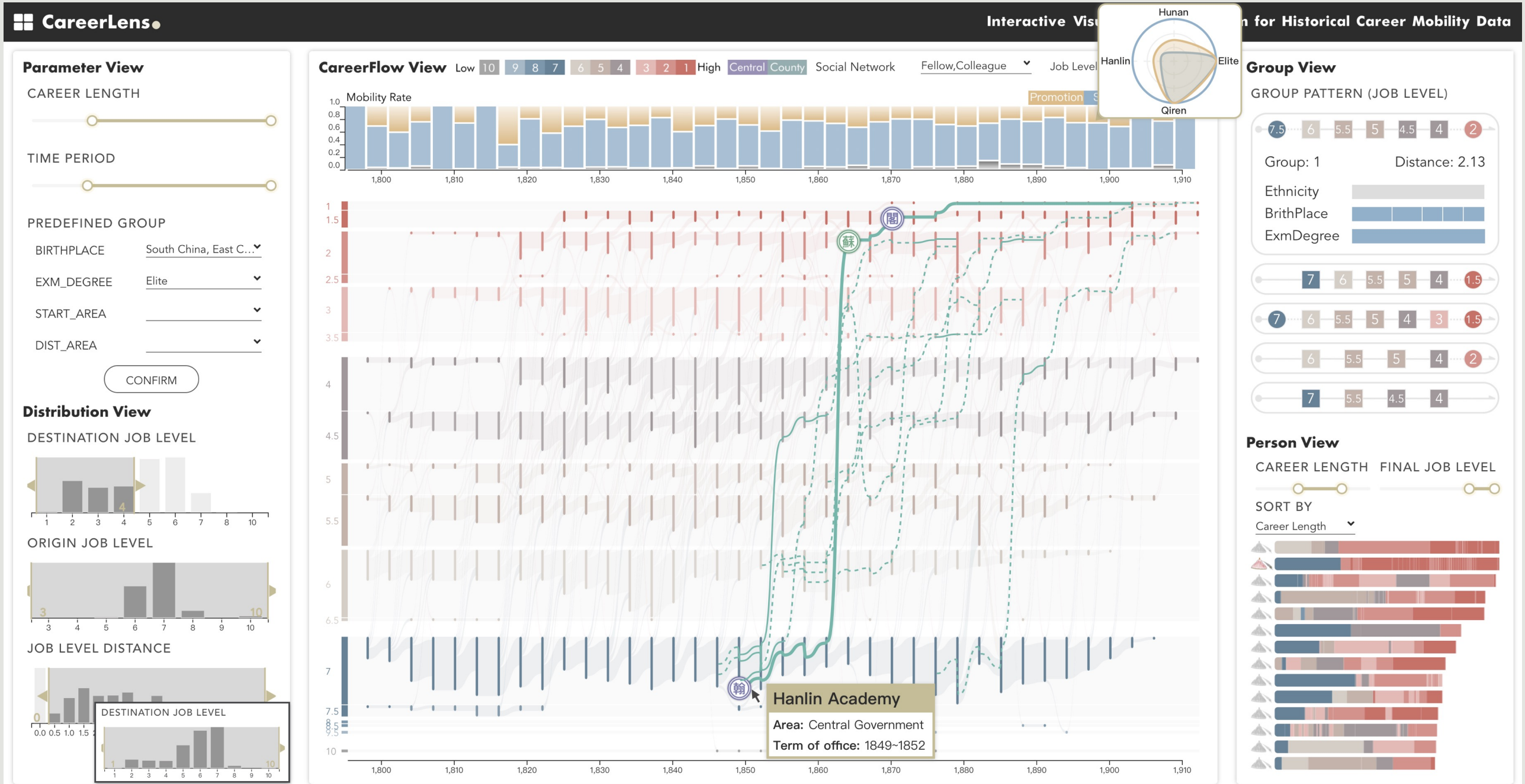
Visual Design

Data Filtering

Visual Design

Visual Design

Detailed Career Mobility Analysis

25
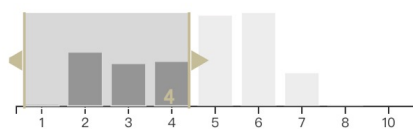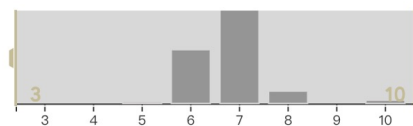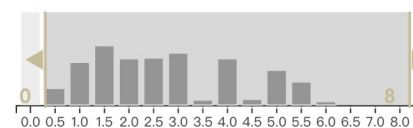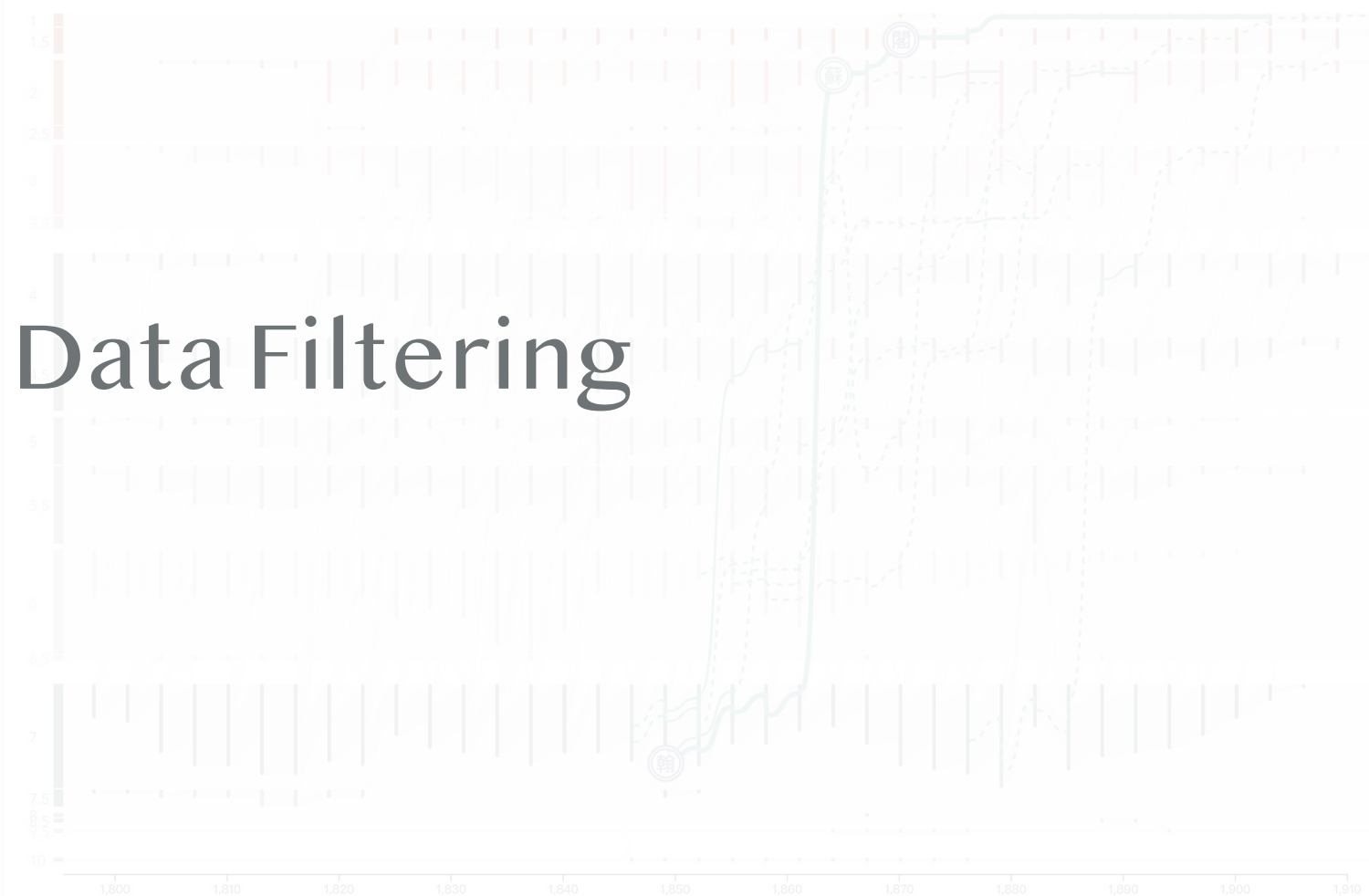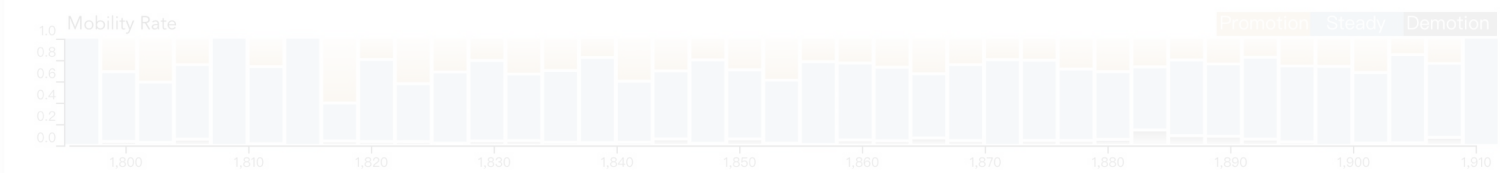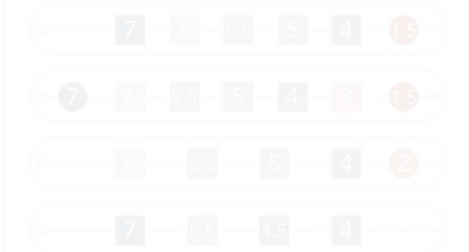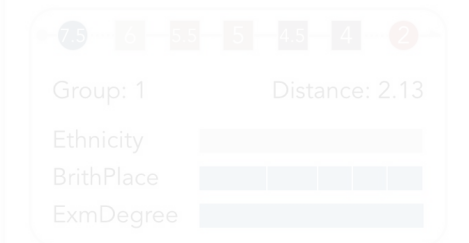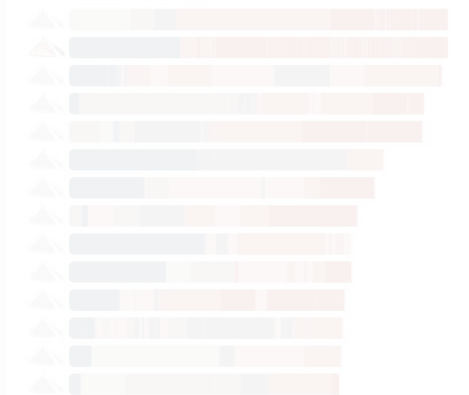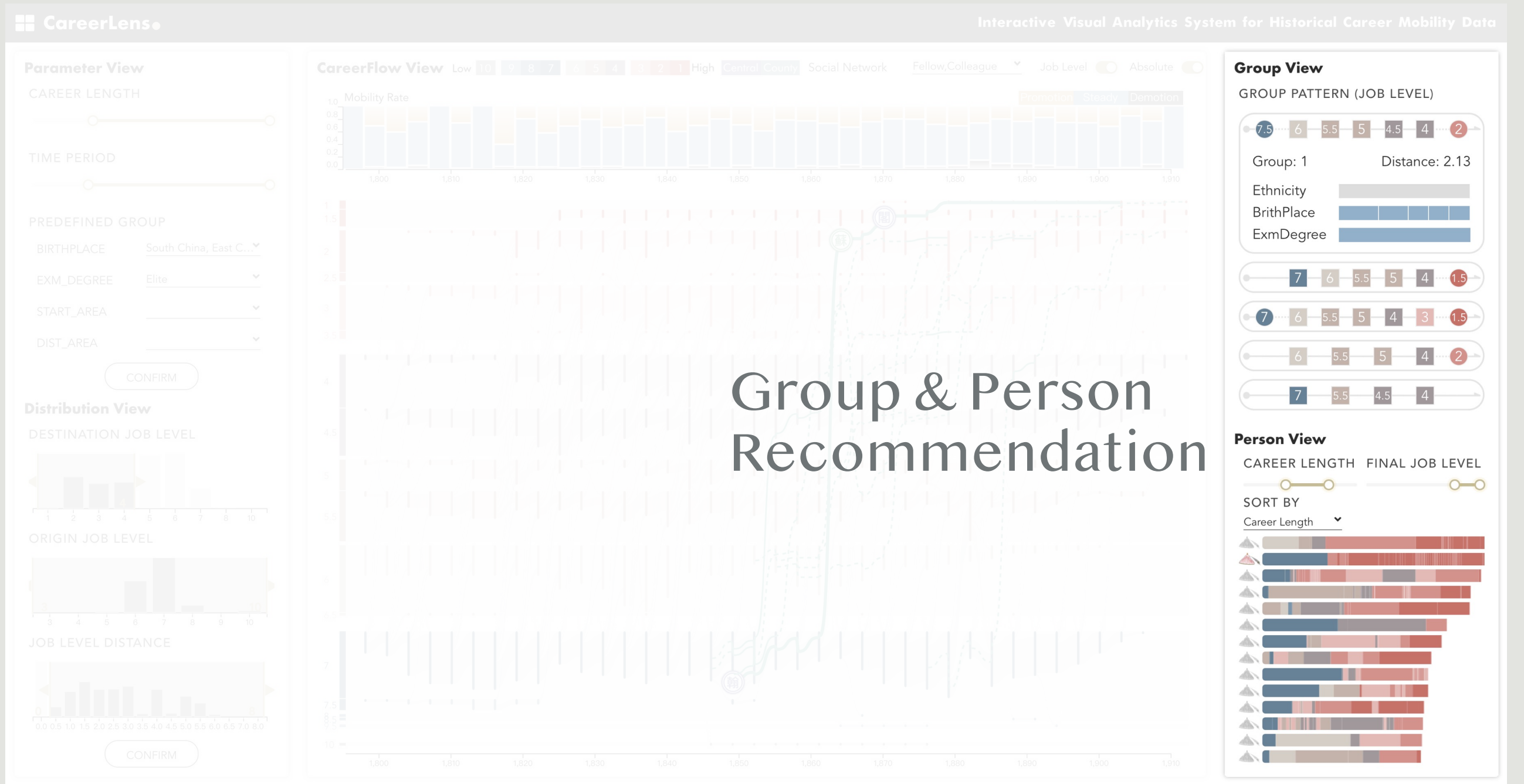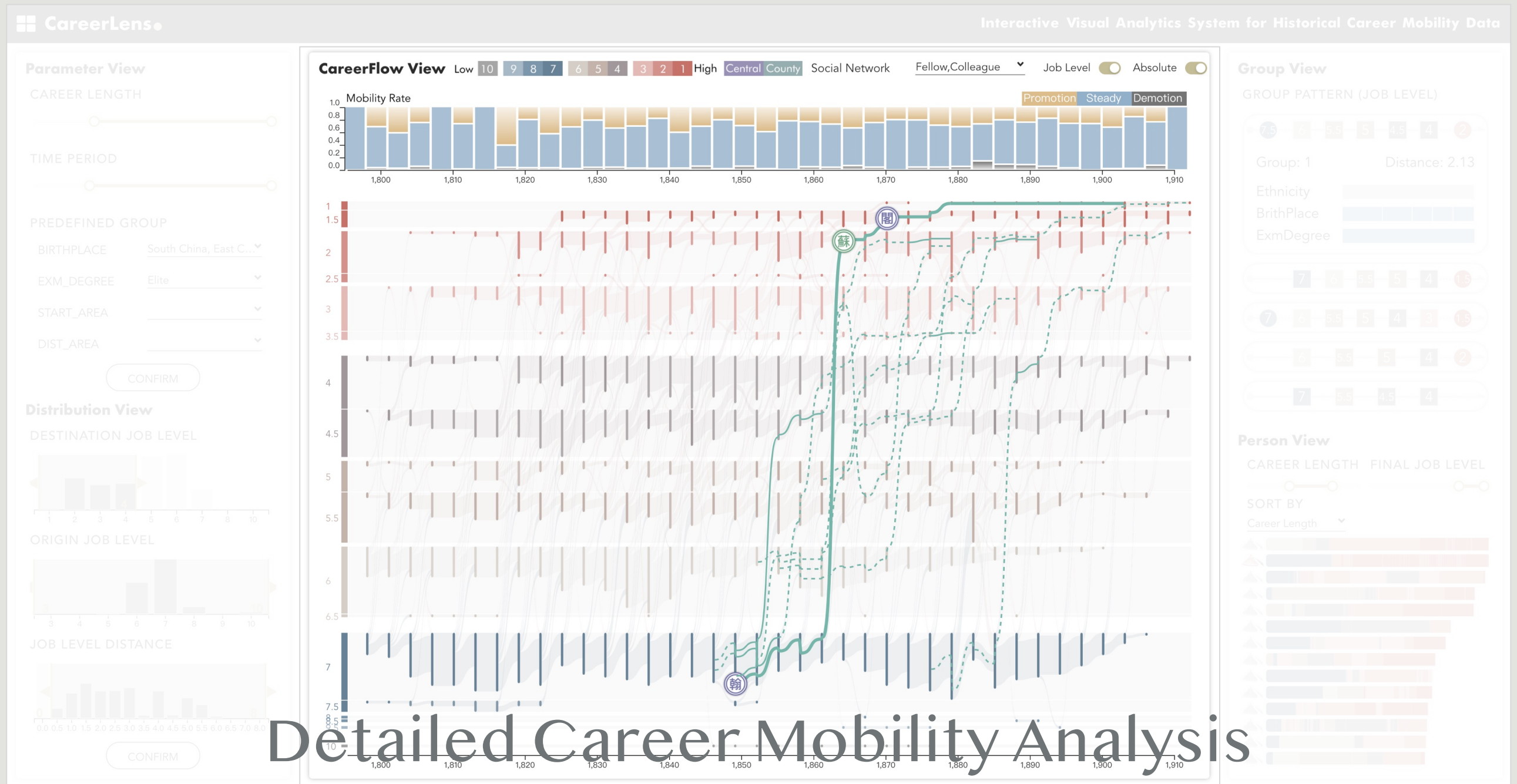
Visual Design

Encoding:
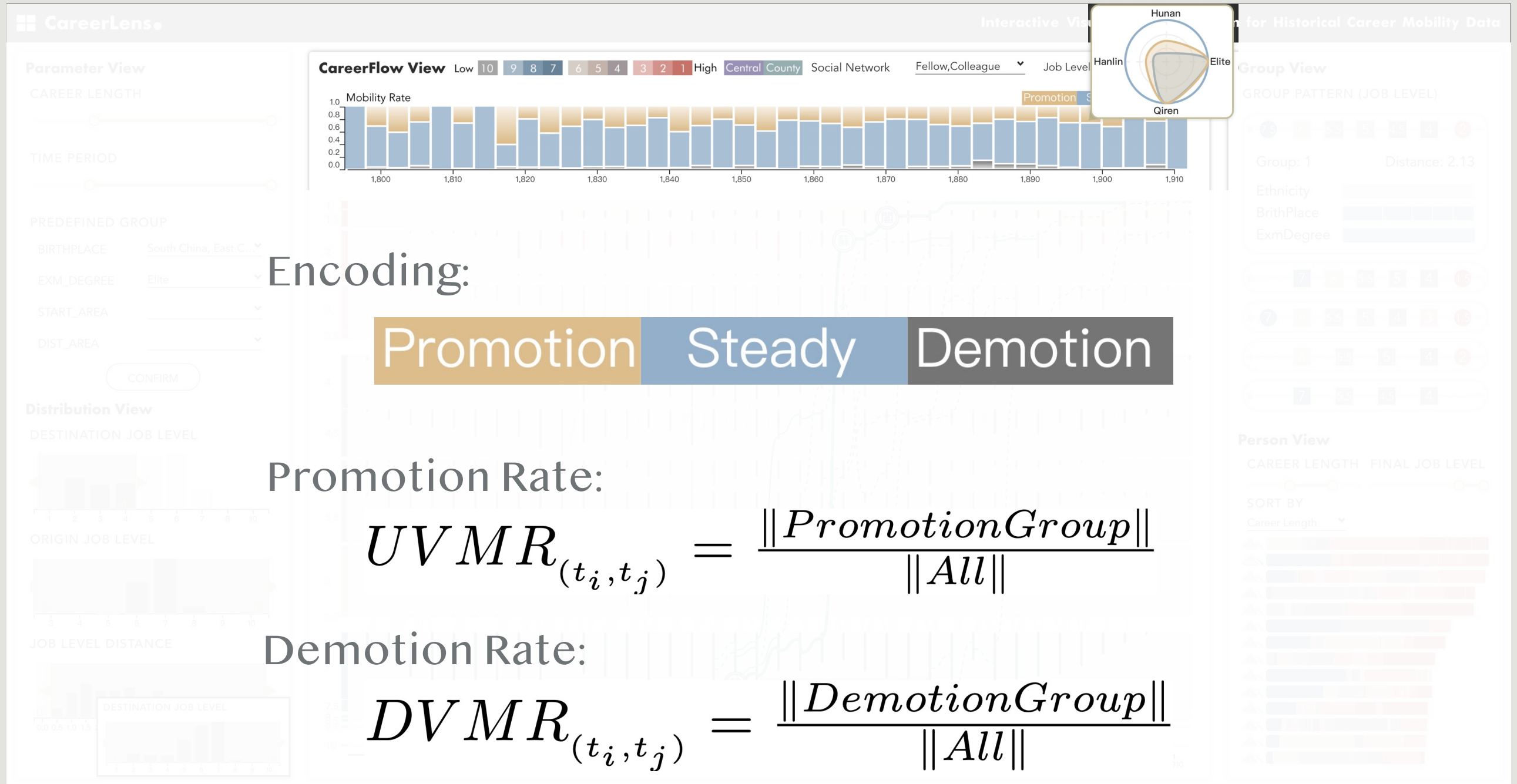
| Promotion | Steady | Demotion |

Promotion Rate:

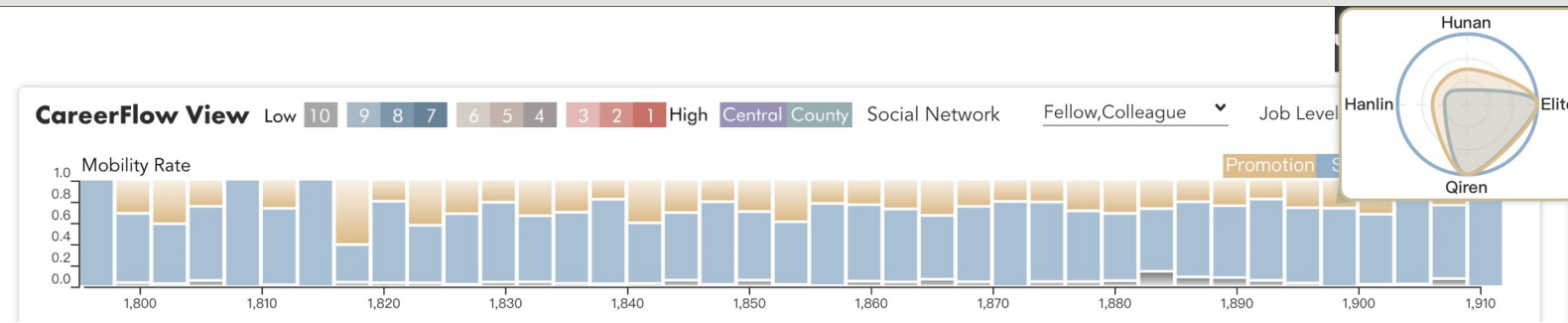$$UVMR_{(t_i, t_j)} = \frac{\|PromotionGroup\|}{\|All\|}$$

Demotion Rate:
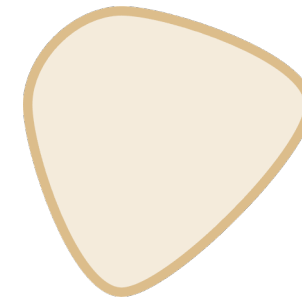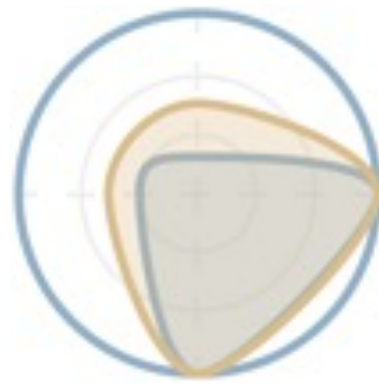
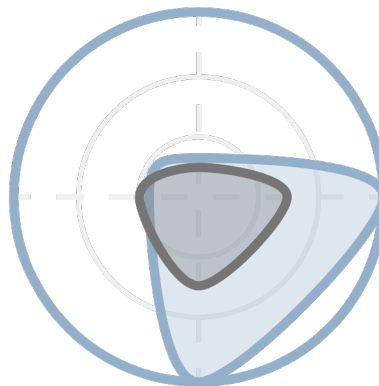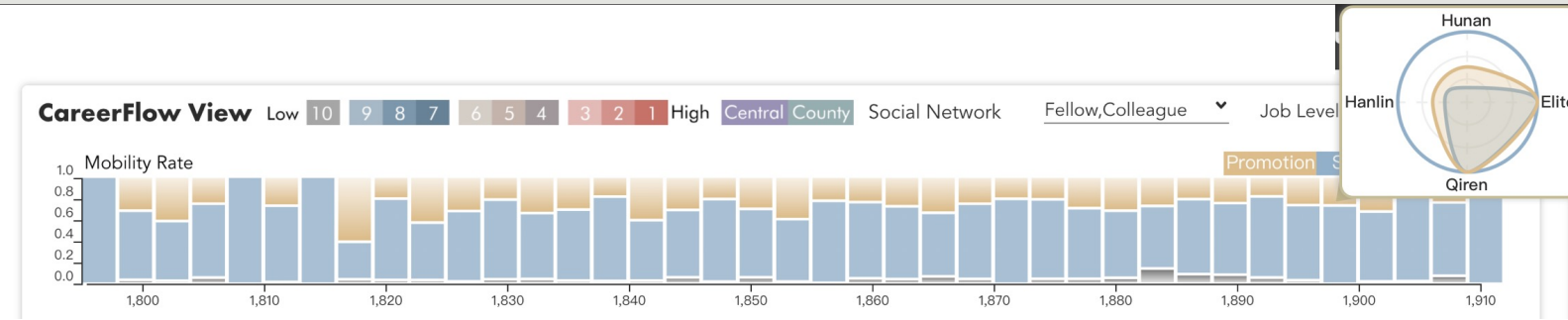$$DVMR_{(t_i, t_j)} = \frac{\|DemotionGroup\|}{\|All\|}$$

Visual Design
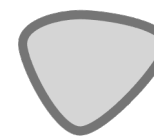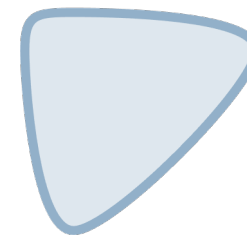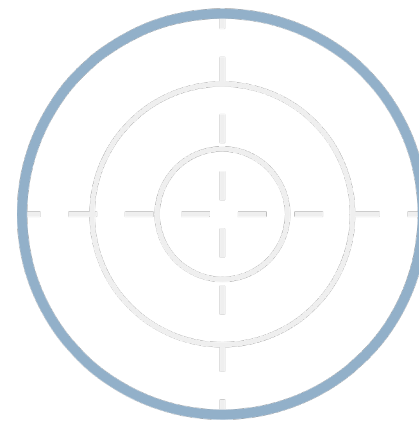
Visual Design

Birthplace
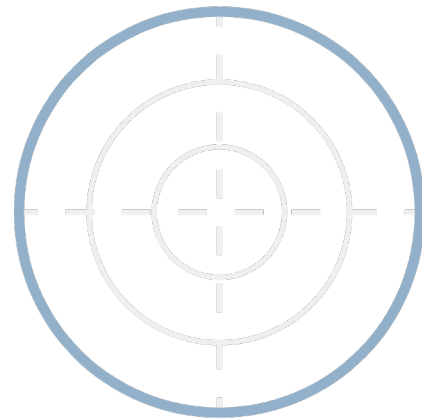
Job
Dept.

Exam
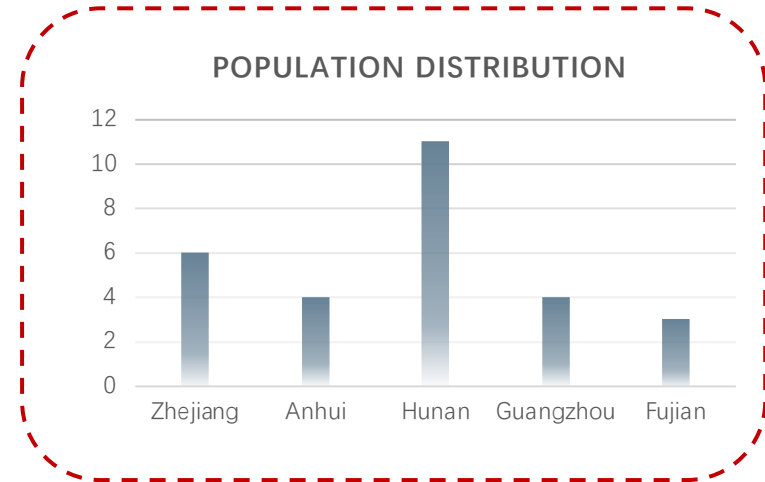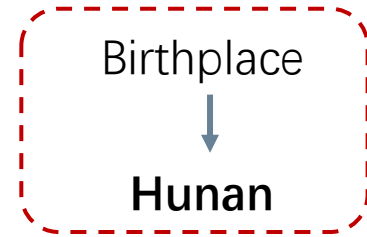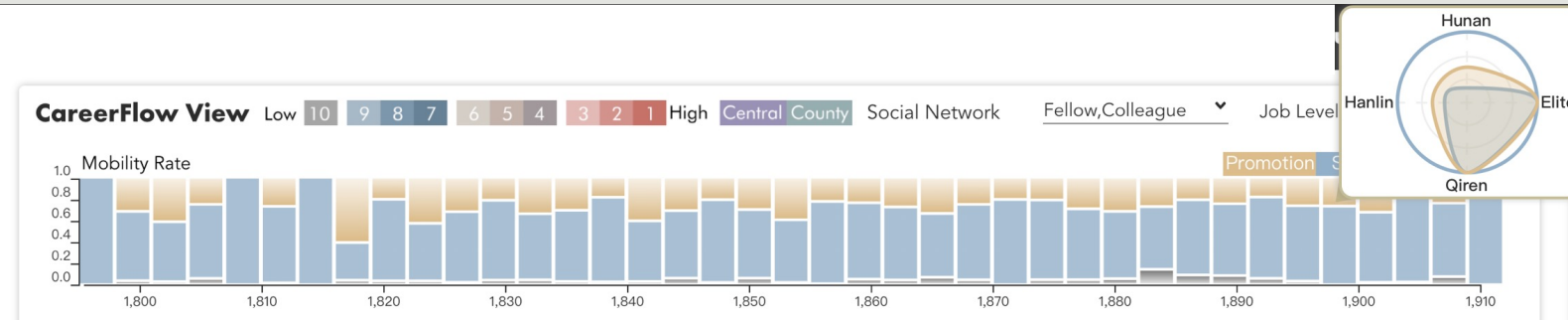Degree

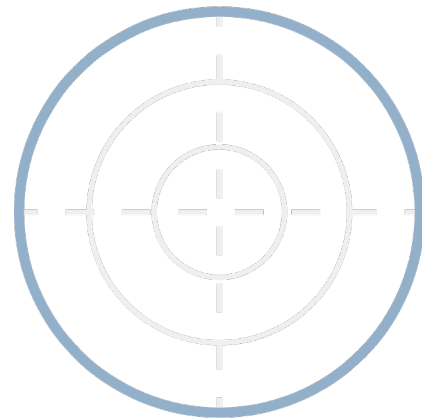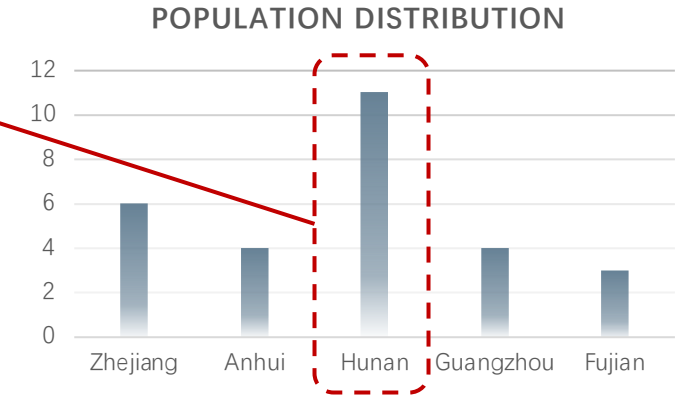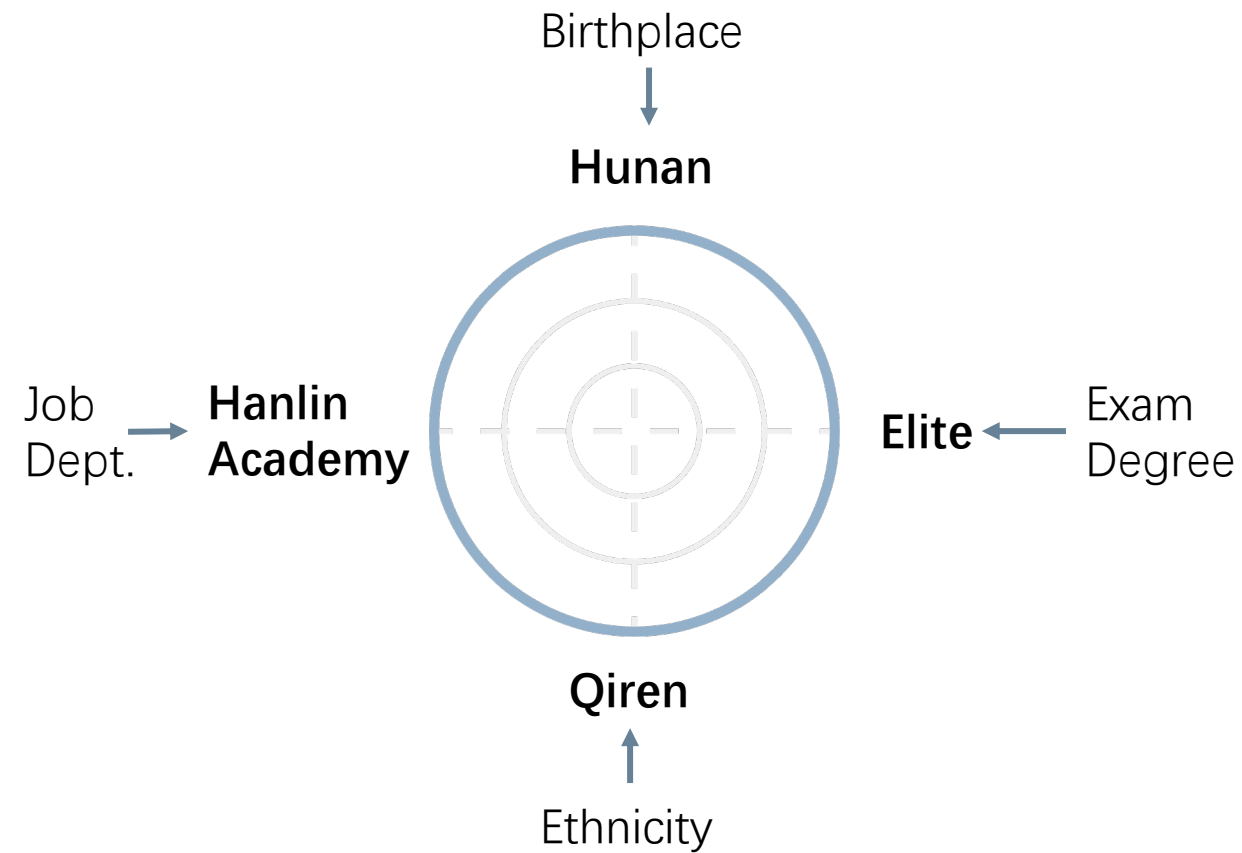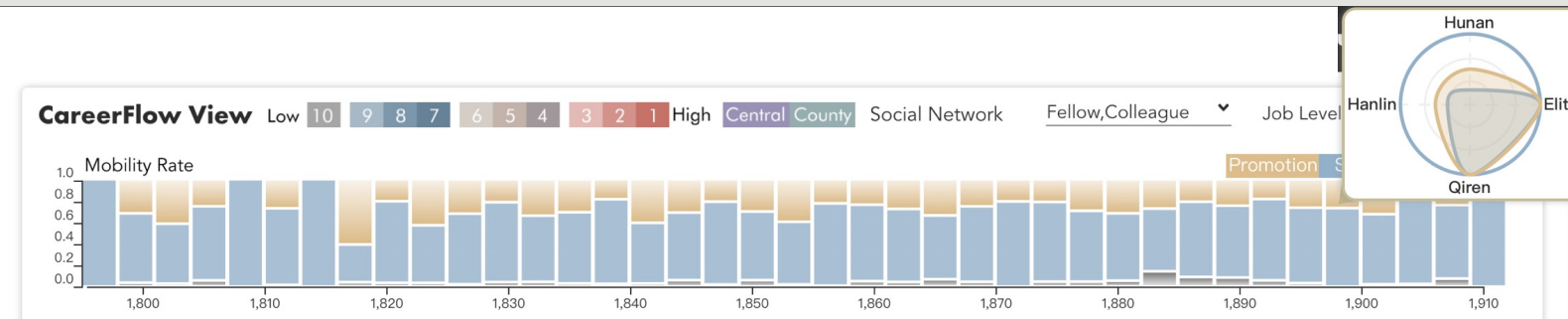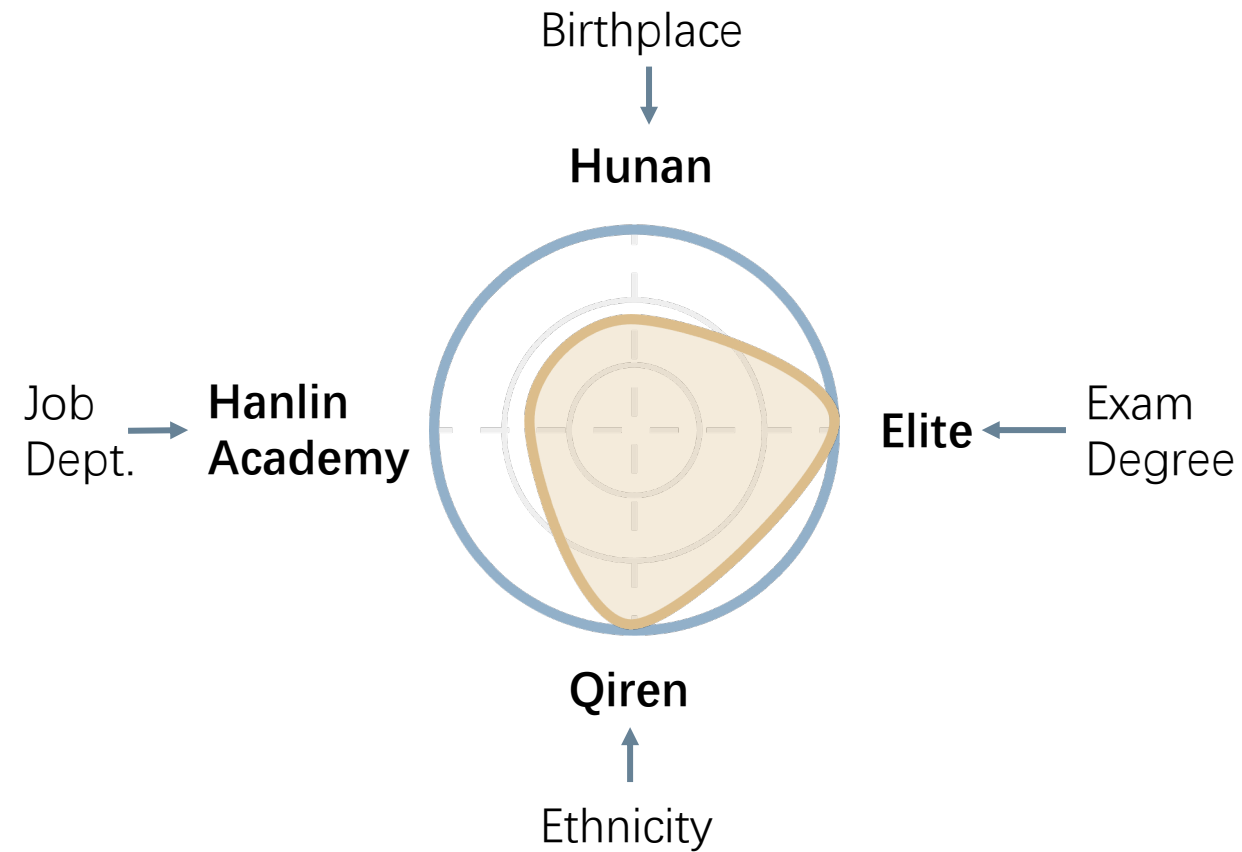Ethnicity

Visual Design

Visual Design

Visual Design
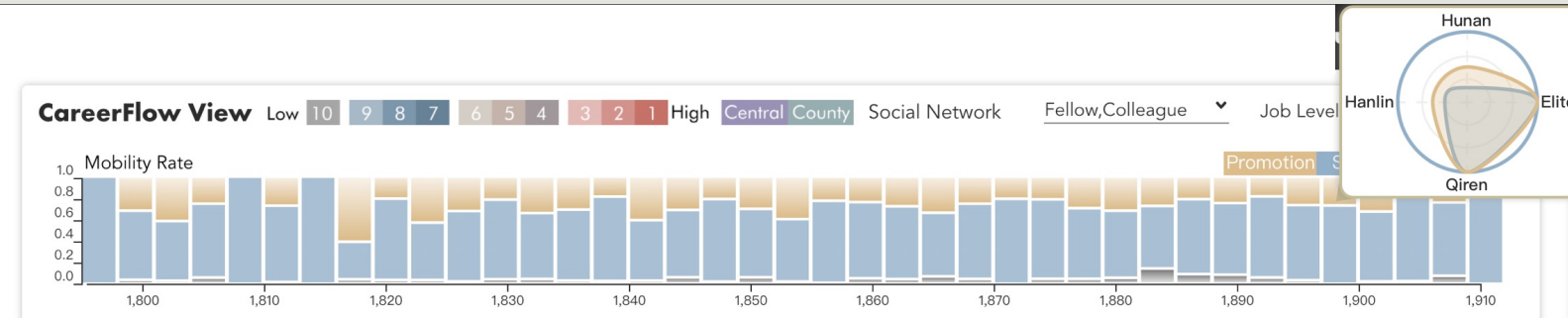
Visual Design
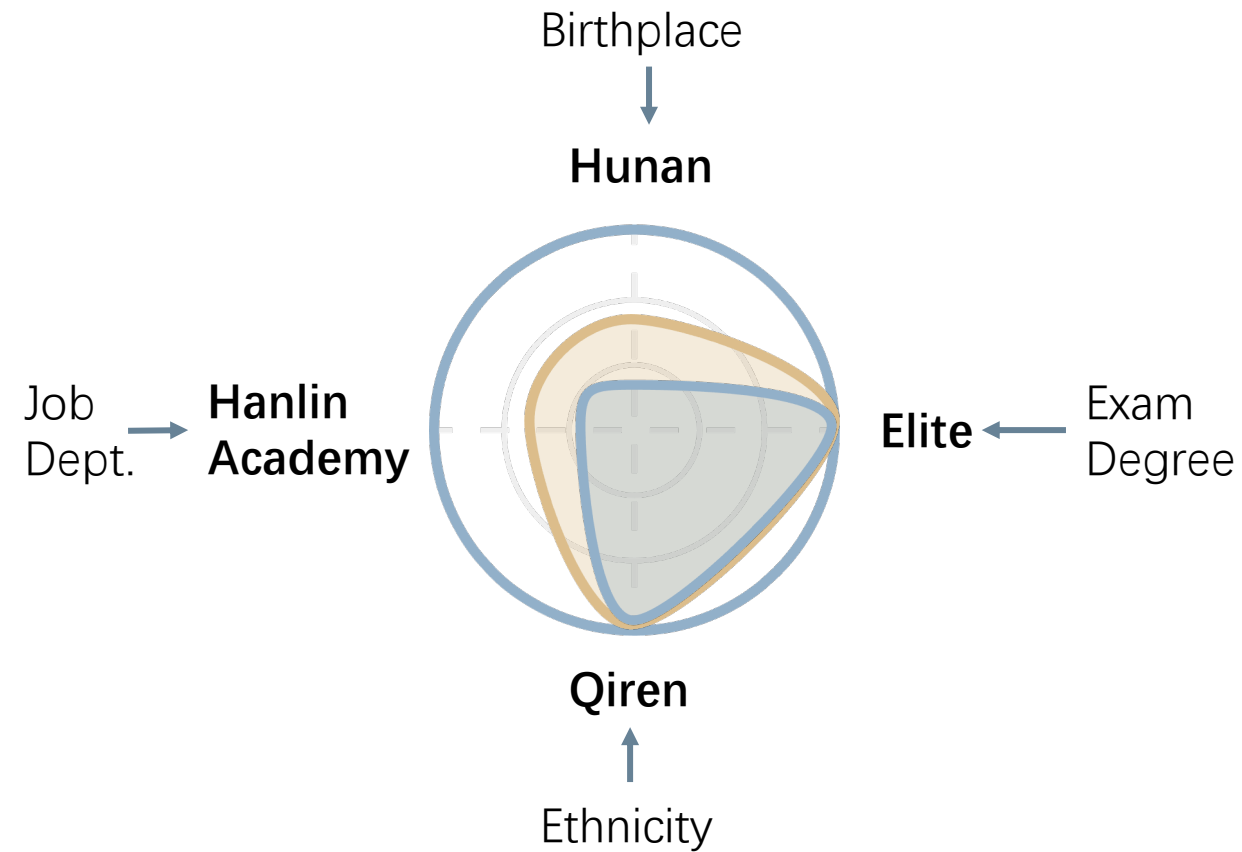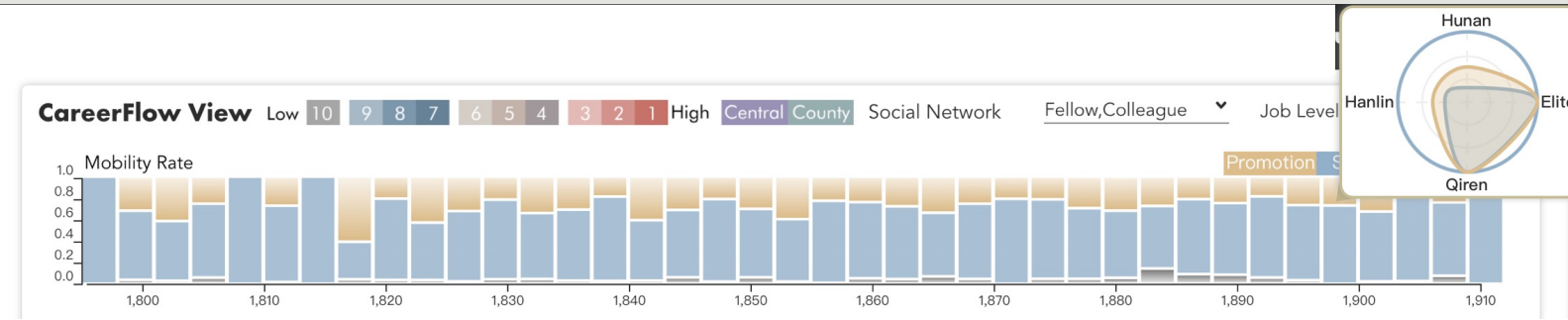
Visual Design

Visual Design

Latent Group Detection (MinDL[1])

$$L(C) = \sum_{(P,G)\in C} \|P\| + \left(\alpha \sum_{(P,G)\in C} \sum_{s\in G} \|edits(s,P)\| \right) + \lambda \|C\|$$

Lowest Job Level   Group Pattern   Highest Job Level
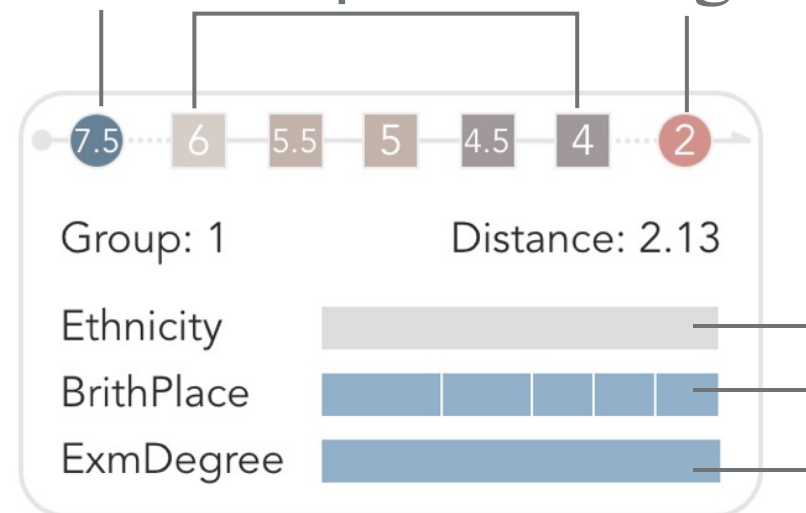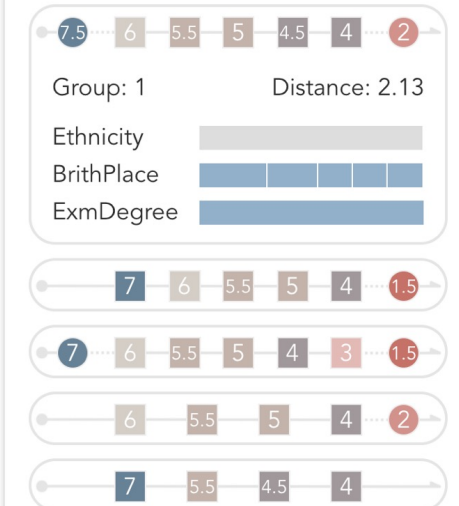
[1] Chen Y, Xu P, Ren L. Sequence synopsis: Optimize visual summary of temporal event data[J]. IEEE transactions on visualization and computer graphics, 2017, 24(1): 45-55.
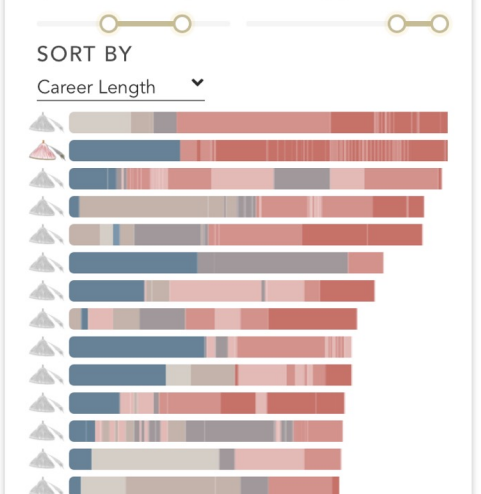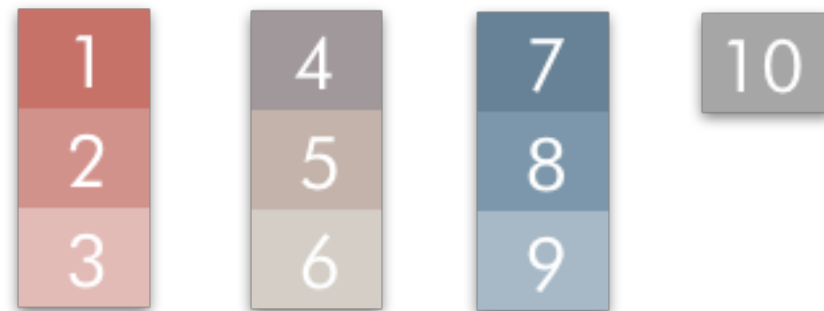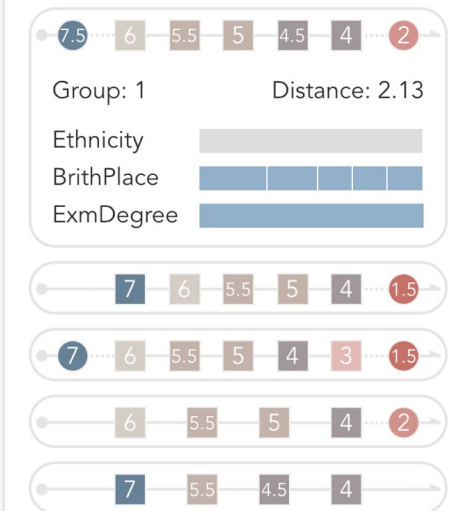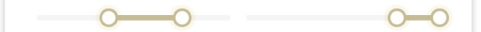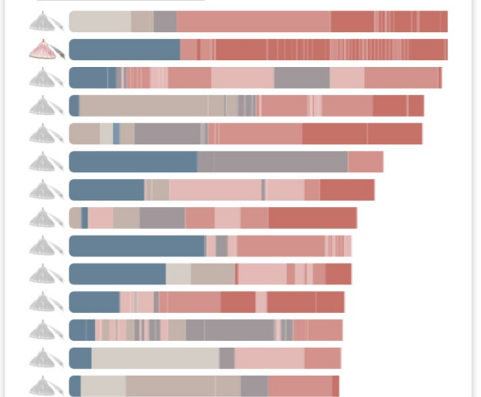
Visual Design

Visual Design

Detailed Career Mobility Analysis (2/2)

Visual Design
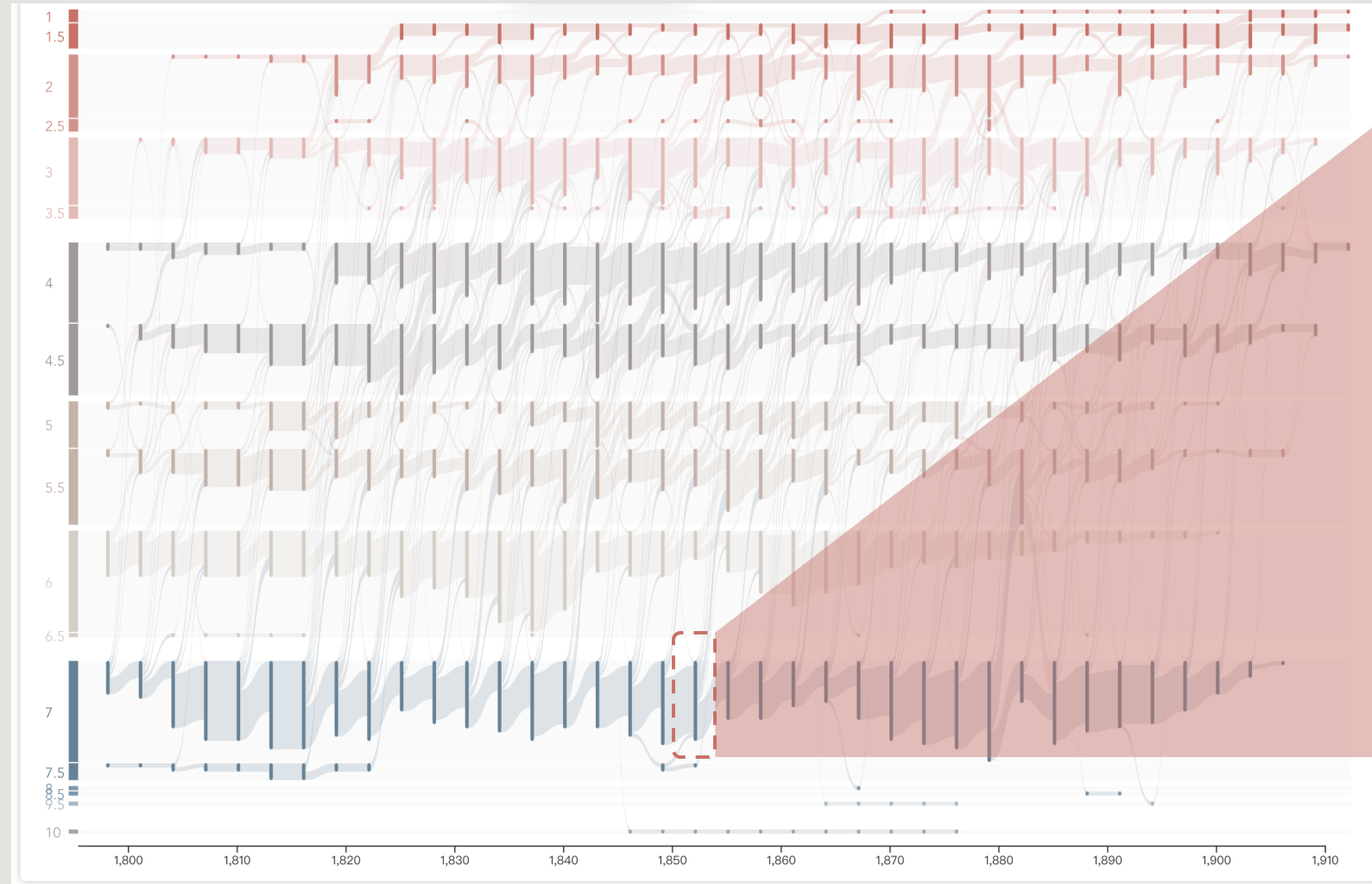
Job-level Mode

Visual Design
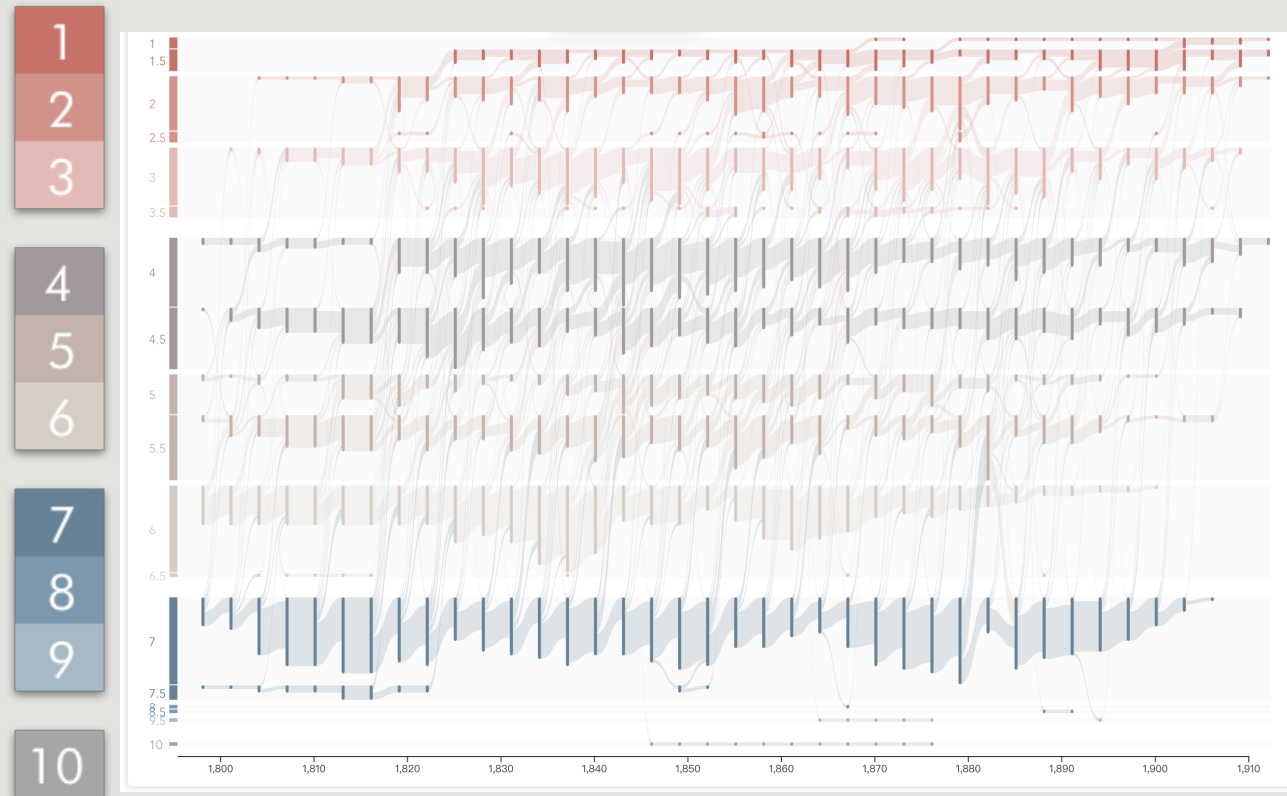
Job-level Mode
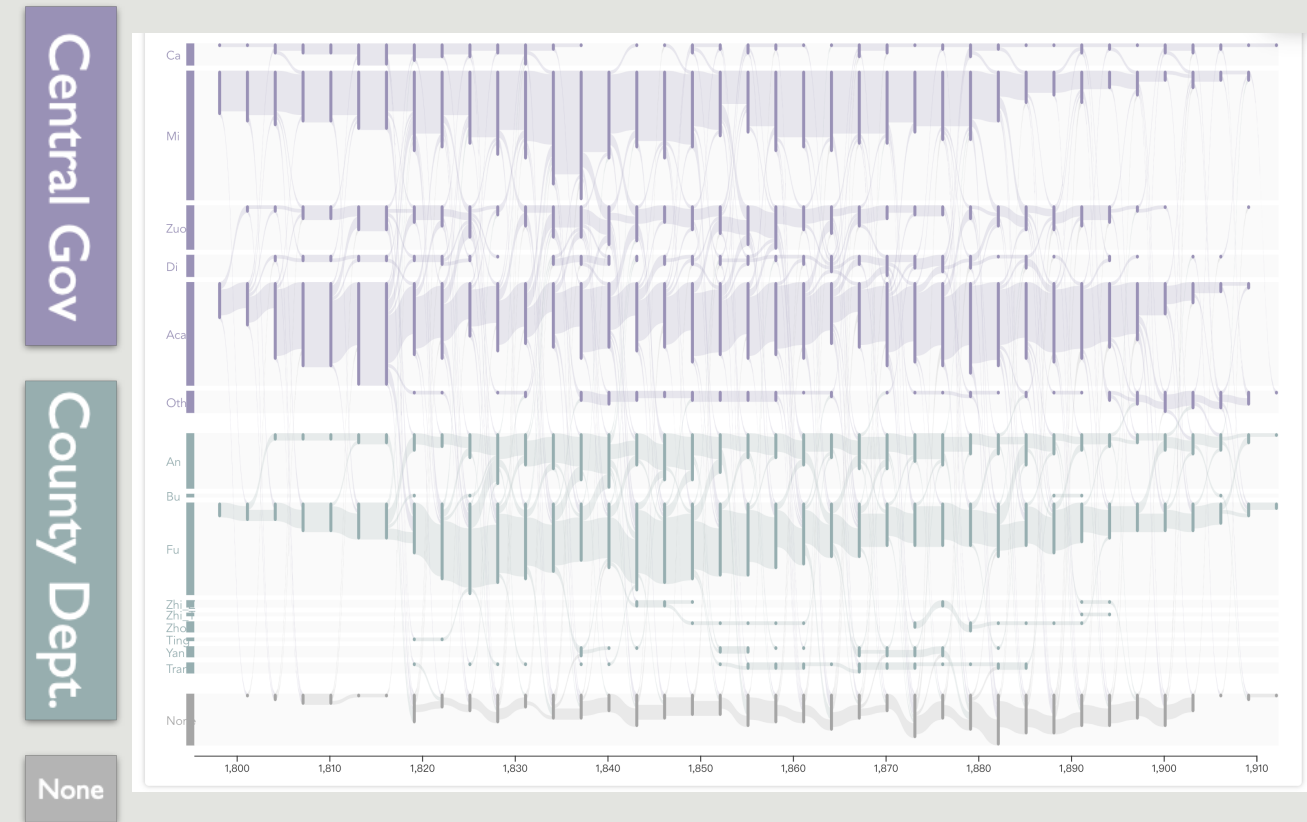
Department Mode

- Color Scheme

40
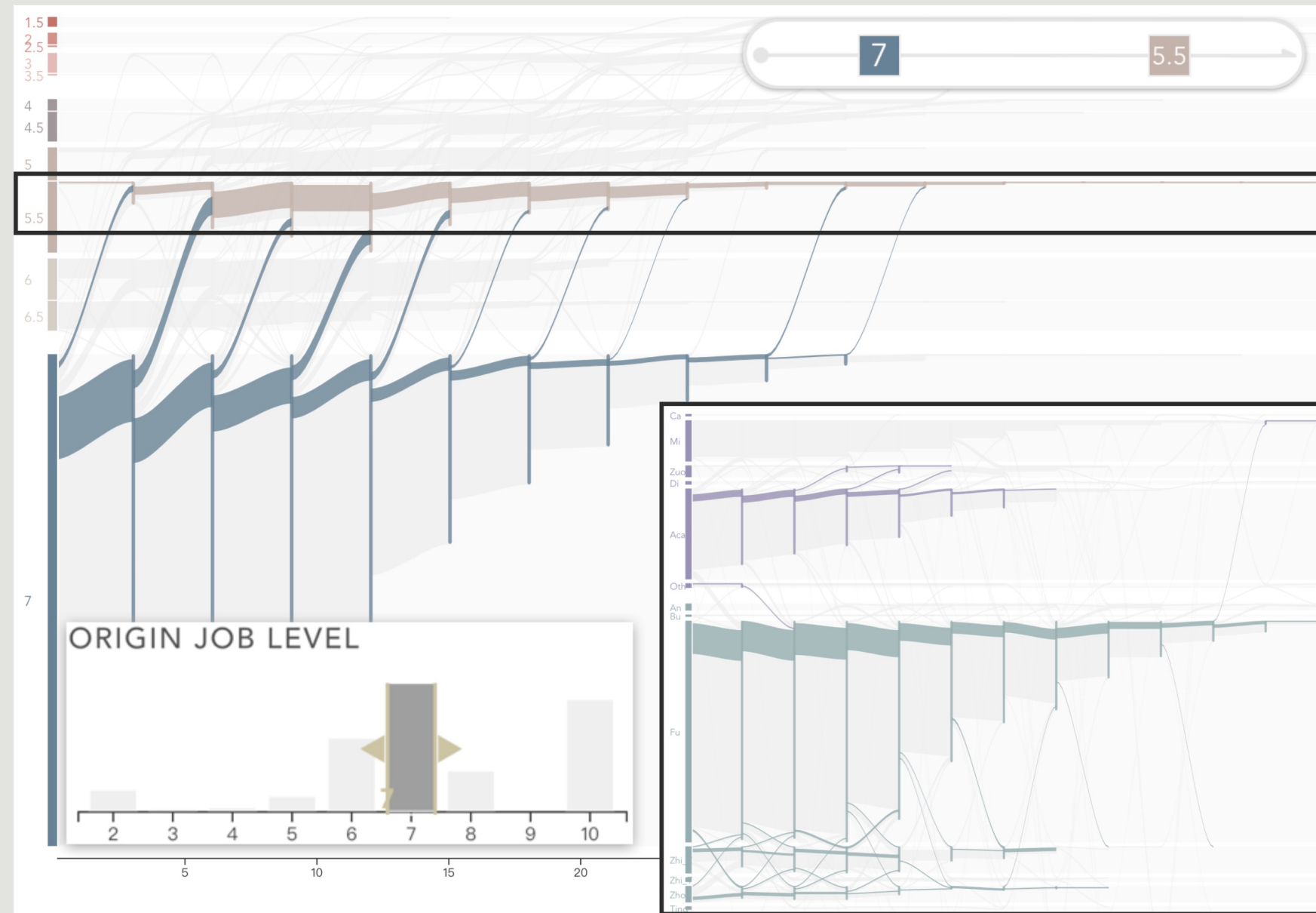
Visual Design

Relative-time Mode: align the career starting year

Visual Design

- Flow Design
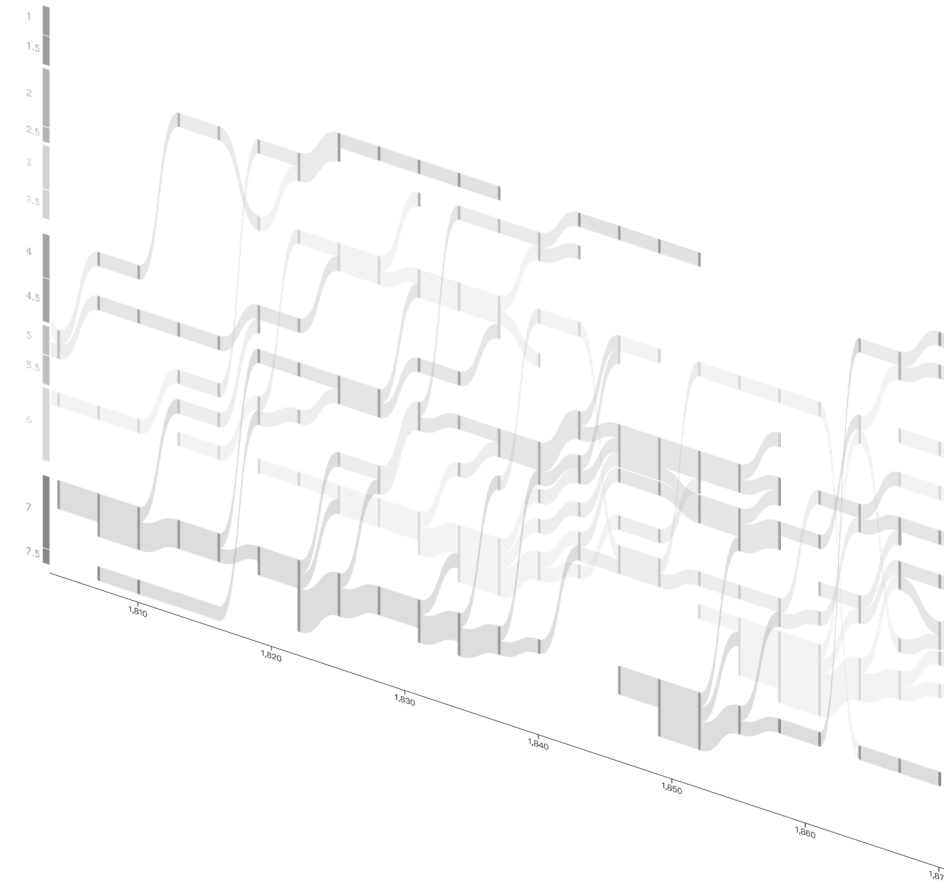
- Overall Flow

Visual Design

- Flow Design

- Overall Flow

- Group Subflow
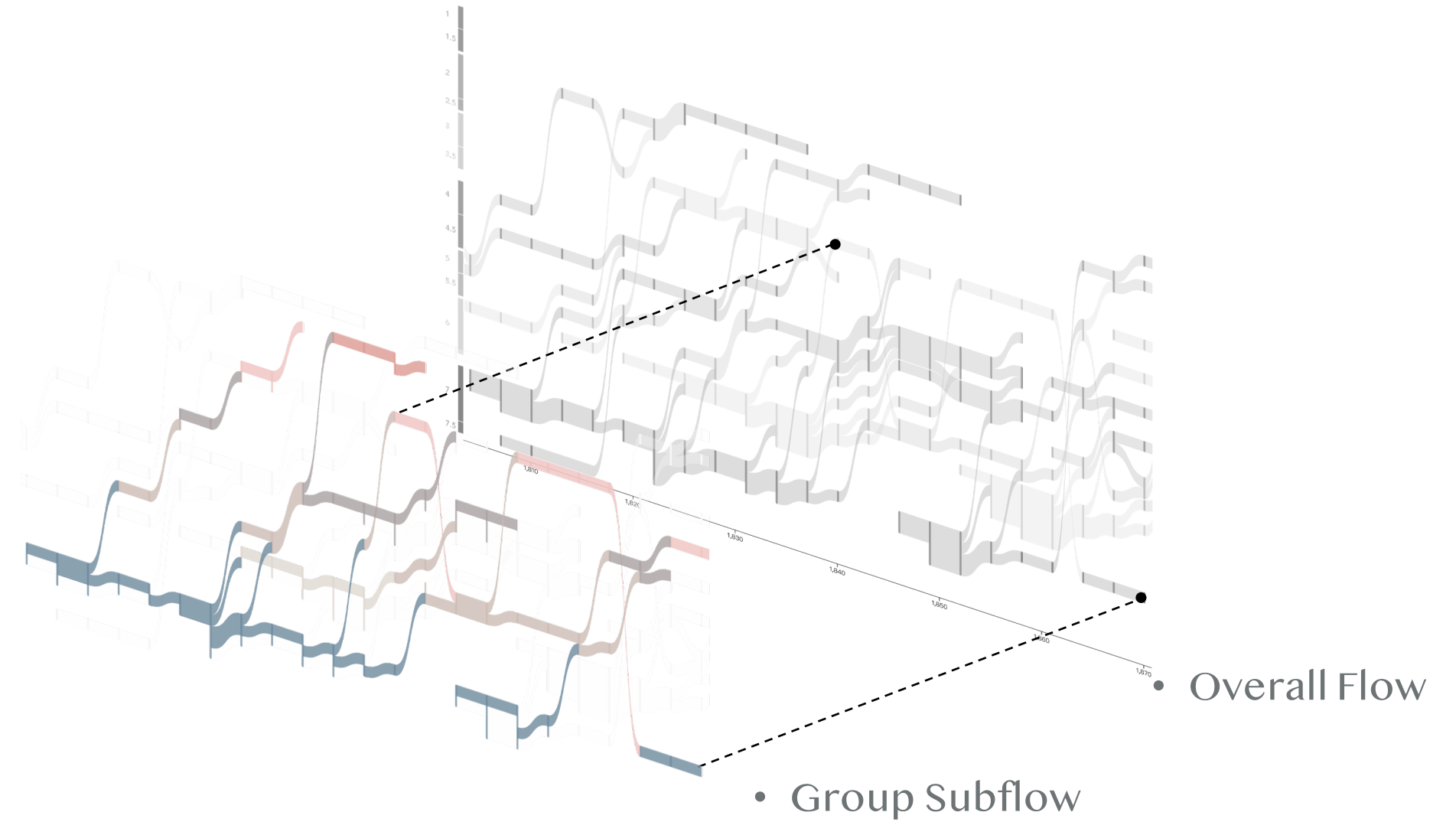
Visual Design

- Flow Design
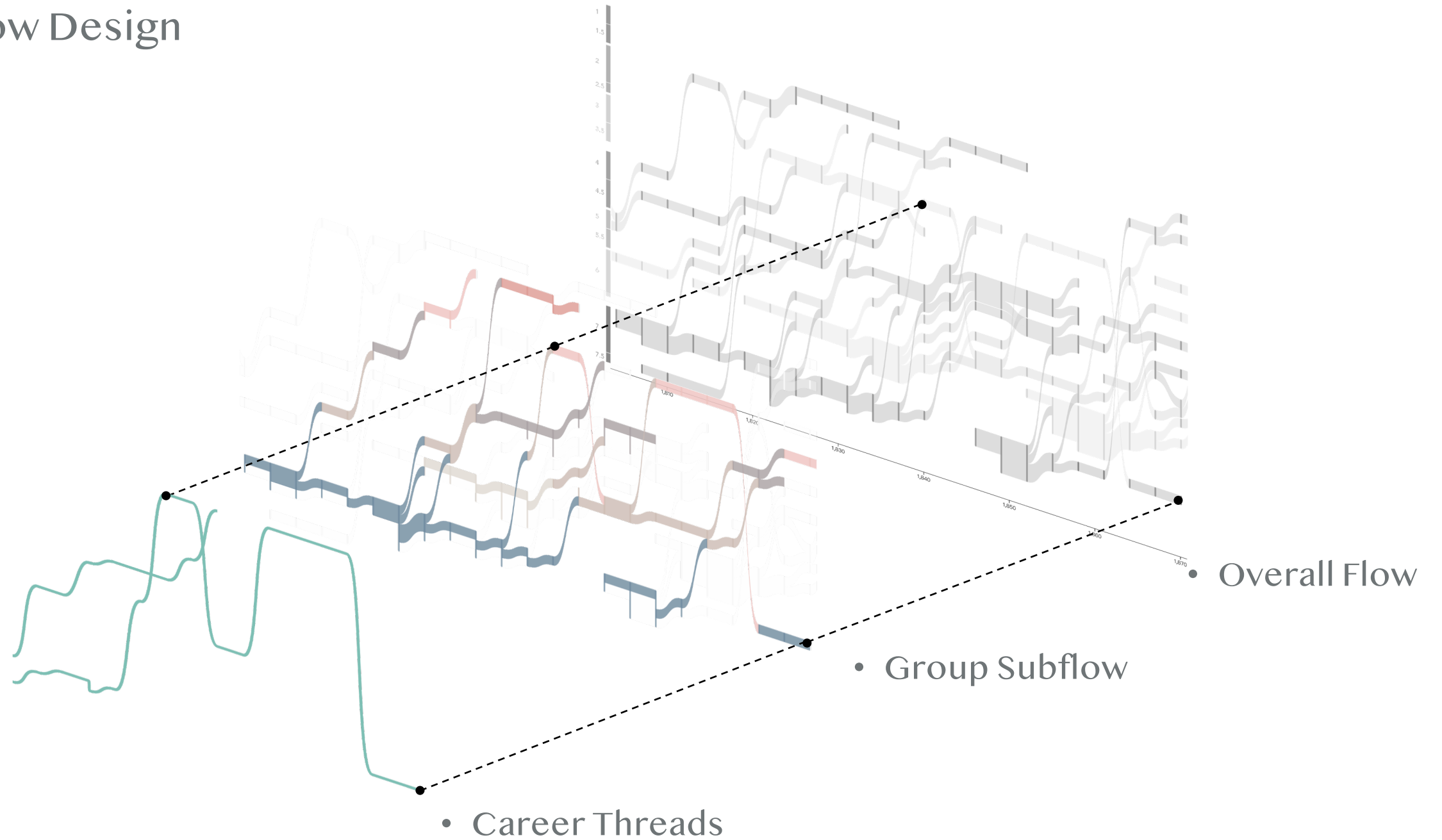
- Overall Flow
- Group Subflow
- Career Threads
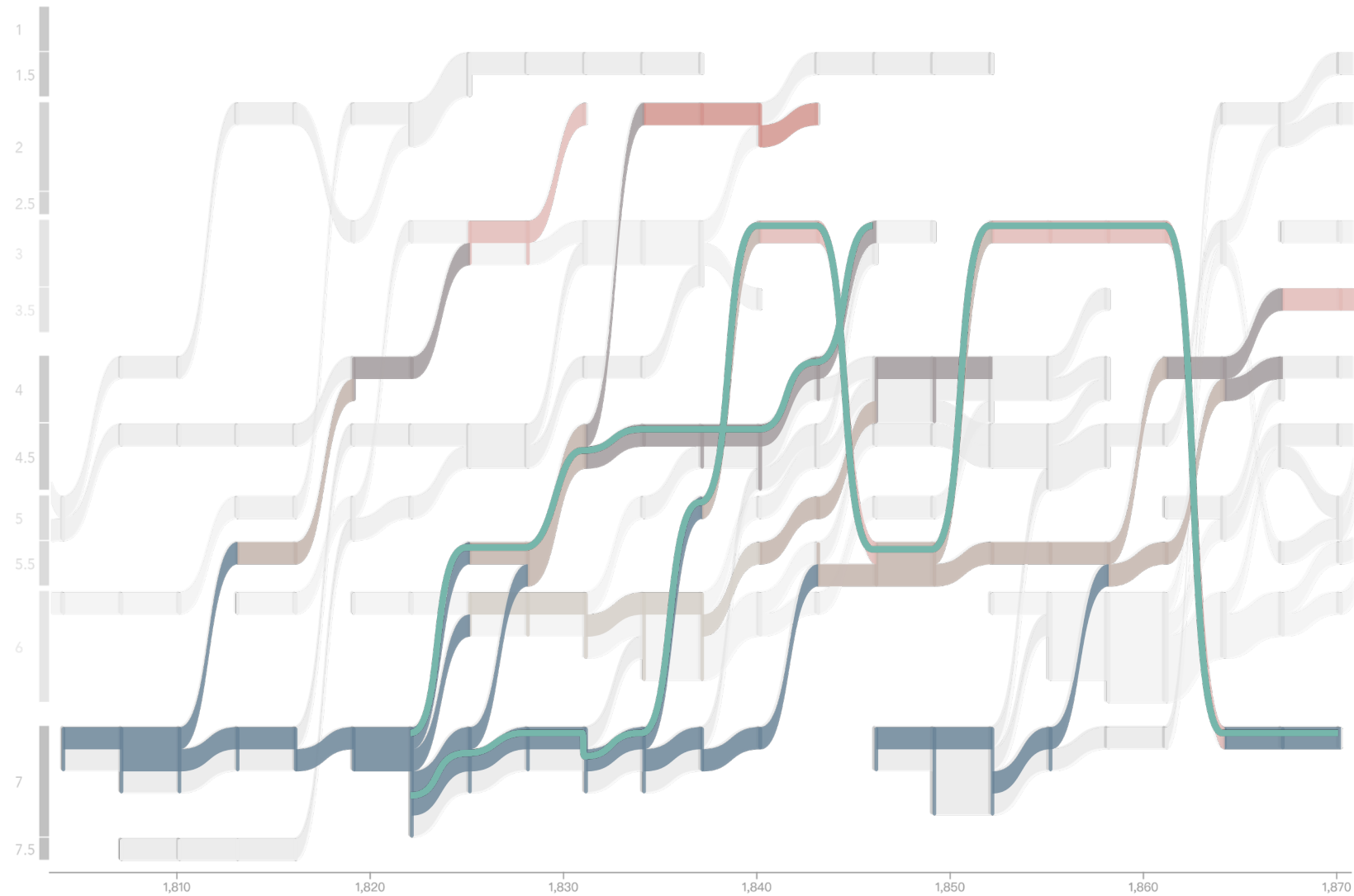
Visual Design
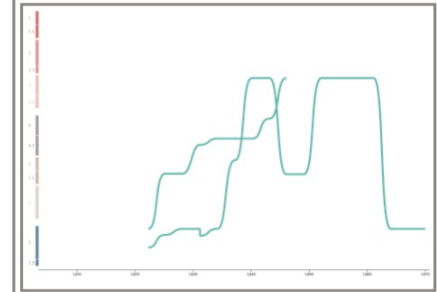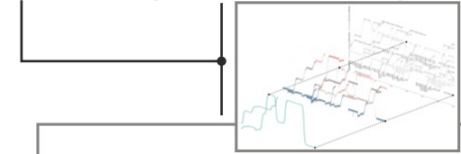
- Flow Design

Career Thread
Highlight several particular individuals
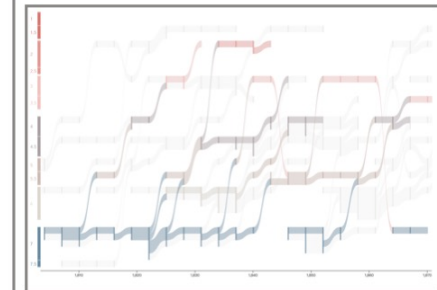
Group Subflow
Aggregate a group into a subflow
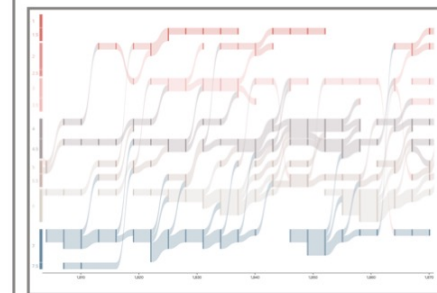
Mobility Flow
Provide a context as overall mobility

Visual Design

- Flow Design + Social Relations of the Key Person (KP)
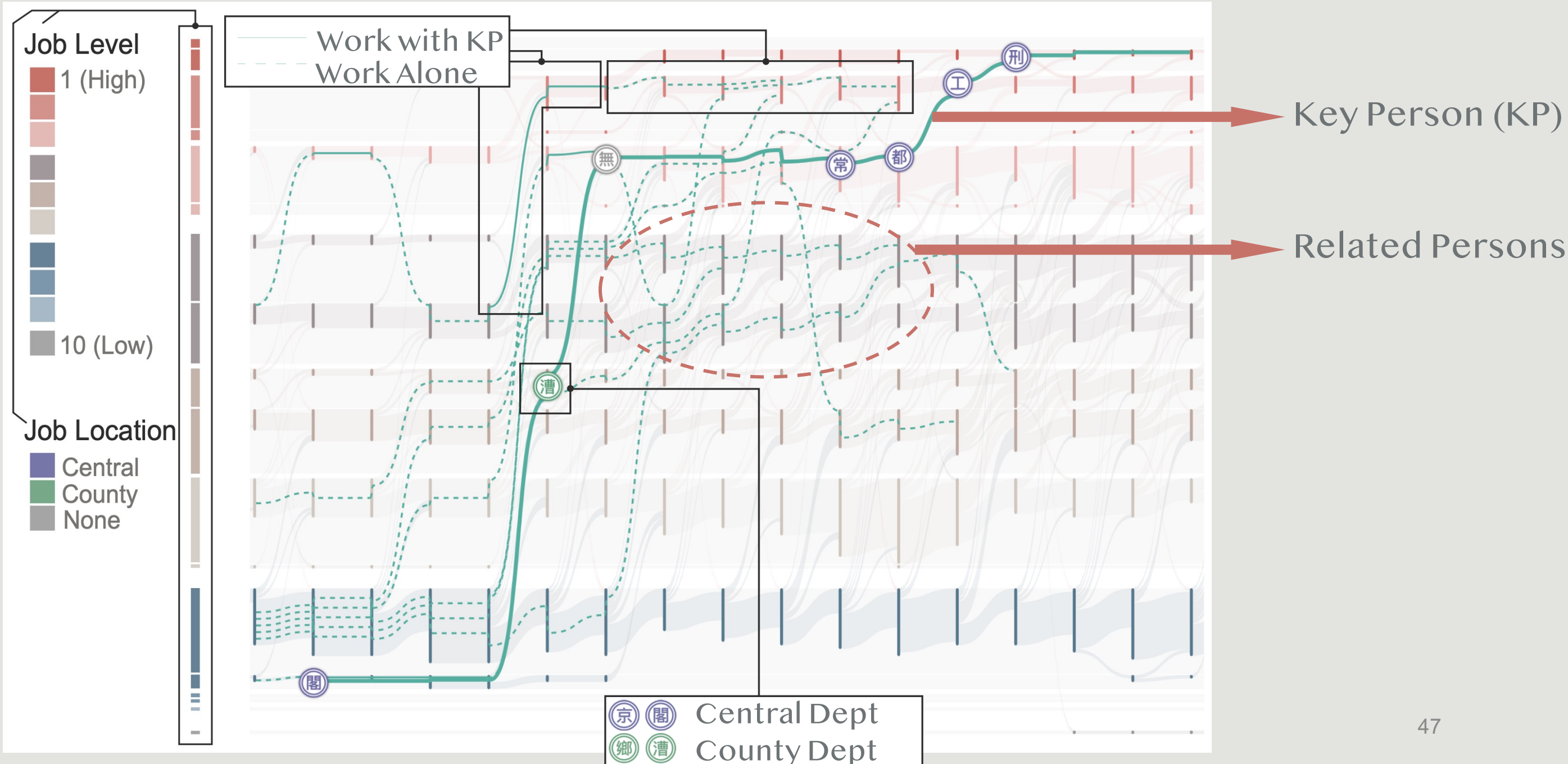
Evaluation

Case Study

- Invite four internal experts to freely explore the system
- Political Elites (i.e., 进士) from South and East China

In this video, we demonstrate a case study to explore the political elites from South and East China.

Expert Interview

- Invite <u>four</u> internal experts and <u>four</u> new experts

- Advantages

  - More efficient workflow to explore the dataset from different LODs with a user-friendly interface
  - Advanced techniques (e.g., latent group and dynamic network) to find interesting insights
  - Intuitive visual representations and interactions

- Suggestions

  - More detailed information
    - Add the departments with provinces
    - Add raw data table
  - Replacing the context menu with a switch button

Conclusion
Future Work
Acknowledgement
Q&A

Conclusion

Conclusions

💡 Contributions
- A hierarchical problem characterization of historical career mobility analysis
- A visual analytics system *CareerLens*
- A novel flow design with a multi-scale approach (overall mobility flow, group subflow, and individual career threads)
- Two case studies and a longitudinal investigation to evaluate the system and receive positive feedback

📝 Future Works
- Open to the public
- Multi-attributes for latent group detection
- Improvement of scalability of flow design
- Correlation of career mobility and different events

For More Information

- I'm currently looking for collaborators in both humanity and social science to work together! For more information, please visit: wangyifang.top or contact yifang.wang@connect.ust.hk

Interactive Visual Exploration of Longitudinal Historical Career Mobility Data
Yifang Wang, Hongye Liang, Xinhuan Shu, Jiachen Wang, Ke Xu, Zikun Deng, Cameron Campbell, Bijia Chen, Yingcai Wu, and Huamin Qu
In IEEE Transactions on Visualization and Computer Graphics (TVCG 2021). [PDF]
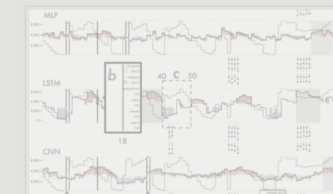
mTSeer: Interactive Visual Exploration of Models on Multivariate Time–series Forecast
Ke Xu, Jun Yuan, Yifang Wang, Claudio Silva, Enrico Bertini
In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2021). [PDF]

CloudDet: Interactive Visual Analysis of Anomalous Performances in Cloud Computing Systems
Ke Xu, Yun Wang, Leni Yang, Yifang Wang, Bo Qiao, Si Qin, Yong Xu, Haidong Zhang, Huamin Qu
In IEEE Transactions on Visualization and Computer Graphics (TVCG 2020). [PDF]

Pulse: Toward a Smart Campus by Communicating Real–time Wi–Fi Access Data
Aoyu Wu, Bon Kyung Ku, Furui Cheng, Xinhuan Shu, Abishek Puri, Yifang Wang, Huamin Qu.
[PDF]
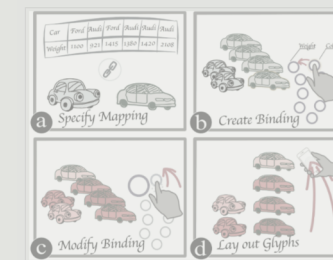
MARVisT: Authoring Glyph–based Visualization in Mobile Augmented Reality
Zhutian Chen, Yijia Su, Yifang Wang, Qianwen Wang, Huamin Qu, Yingcai Wu.
In IEEE Transactions on Visualization and Computer Graphics (TVCG 2019).
Doi: 10.1109/TVCG.2019.2892415. [PDF]
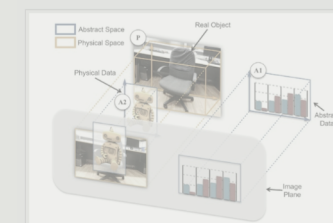
Exploring the design space of immersive urban analytics
Zhutian Chen, Yifang Wang, Tianchen Sun, Xiang Gao, Wei Chen, Zhigeng Pan, Huamin Qu, and Yingcai Wu.
In Visual Informatics, 1(2):132–142, 2017. [PDF]

Interactive Visual Exploration of Longitudinal Historical Career Mobility Data

Q&A

Yifang Wang[1,2]

Hongye Liang[2,3]

Xinhuan Shu[1]

Jiachen Wang[2,3]

Ke Xu[4]

Zikun Deng[2,3]

Cameron Campbell[1]

Bijia Chen[5]

Yingcai Wu[2,3]

Huamin Qu[1]

- I'm currently looking for collaborators in both social science and humanities to work together!

- For more information, please visit: wangyifang.top or contact yifang.wang@connect.ust.hk

1 香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

2 浙江大学
ZHEJIANG UNIVERSITY

3 之江实验室
ZHEJIANG LAB

4 NYU

5 中国人民大学
RENMIN UNIVERSITY OF CHINA