KIMEP Data Science Lab Meeting Minutes April 15, 2022

14:00-15:00 - Python for Beginners

15:00-16:30 - Main Session:

1. Team Updates

| **No.** | **Team** | **Project** | **Project Update** |
| --- | --- | --- | --- |
| 1 | Education Policy | policy recommendations to promote more equal distribution of educational resources | Last week: considering data sources   * Resources (money, labs) * Demographics: disabled, low-income families   Update: brainstorm indicators   * School description (profile; building; geography; toilet in school; # of classrooms, labs) * Resources (books, computers,) * Stakeholders (admins, teachers, students) * Funds (budget, * Overall demographics (university enrollment, student success)   *Next update:*   * *look for existing data sources* |
| 2 | Consulting Team 1: Magzhan, Azamat, Alan, Aida | industry consulting | Last week: will choose an industry and list 5-10 companies who you would like to approach  Update: |
| 3 | Team TechnoDom: Aruyel, Altynai, Tursunai, Aidana, | Industry consulting | Last week: working on a proposal  Update: → Team PWC |
| 4 | North Korean Elite Governance:  Jacob, Seung-ok, Dilbar, Zhandos | academic publication analyzing internal politics of North Korea using elite resumes | Last week: planning to meet and look at data for the first time; currently updating data  Update: |
| *5* | *South Korean media analysis* | *academic publication using NLP and semantic networks to analyze South Korea’s North Korean Human Rights movement* | *Last week: No updates. An ongoing project open to new collaborators.* |

1. Business Items
   1. **New efforts to solicit consulting jobs through KIMEP alumni**
   2. **Two Sessions**
      1. **2-3pm - Coding for Beginners: Installing software. Basic tutorials. (Nariman)**
      2. **3-4:30pm - Lab Meeting: Discuss projects. Advanced analysis and tutorials**
2. [Introduce North Korean elite resumes](https://nkinfo.unikorea.go.kr/nkp/theme/peopleList.do)

MOUID, KoreanName, BirthString, DeathString, BirthYear, BirthCountry, BirthCP, BirthBCC

MOUID, KoreanName, CareerStringID, CareerString, CareerStartYear, CareerStartMonth, EventType, JointAppointment, Org1, Org2, Org3, Org4, Org5, Position, OrgLevel, PositionLevel, OrgPositionString

1. Pandas in North Korea - using Pandas to resolve some issues

* Joining, merging data
* Matching strings
* Creating data subsets

**Notes and Links for Python, Jupyter Notebooks and Pandas**

Installing Python, Jupyter Notebook, Pandas

* Python - <https://www.python.org/>
* Jupyter Notebook - <https://jupyter.org/>
* Python Pandas - <https://pandas.pydata.org/>

Pandas Cheat Sheets

* <https://towardsdatascience.com/my-python-pandas-cheat-sheet-746b11e44368>
* <https://www.dataquest.io/blog/pandas-cheat-sheet/>

Topics in Pandas

* Import/export data
* View
* Clean
* Panda & SQL
  + Select
  + Filter
  + Group
  + Order
  + Join/Combine
  + Statsitics