|  |  |
| --- | --- |
| Data Analyst/ Business Analyst  Arman nurken | +7 705 232 07 41 | arman.nurken04@gmail.com  <https://www.linkedin.com/in/rmanotheanalyst/>  Portfolio Website: <https://rrmanoo.github.io/RmanoTheAnalyst.github.io/>  Astana, 010017 |

# Skills

* SQL (SQL Server, MySQL, PostgreSQL) & NoSQL
* Python (Pandas, NumPy, SciPy, MatPlotLib)
* Tableau
* Excel (X-VLookup, Conditional Formatting, Pivot Tables)
* Microsoft Power BI

# Projects

COVID-19 June 2023

* Utilized **SQL** to explore global COVID infection data: death percentage, population infected per country and continent, forecast to significant countries.
* On **Tableau** visualization used **Map, Line graph and bar charts**.

Breakdown to data field Survey June 2023

* Transforming data in **Power BI**: Splitted columns with inaccurate data, instead left option “other”. Reformatted salary range to average for visualization.
* Visualizations: **Card** (total surveyors, avg age), **Bar chart** (avg salary by job titles), **Column chart** (Favorite prog. Lang.s by amount of voters), **Treemap** (specifying countries), **Gauges** (Happiness with salary and work & life balance), **Donate** (difficulty to Break into Data)

Movie correlation June 2023

* Utilized **Python** for analysis: dropped rows with the missing and duplicate data, sorted by gross earnings
* **Scatter plot** to Budget vs Gross and applied **linear regression line**.
* Made **Hypothesis**: Gross affected by budget and company
* Plotted **correlation matrix** by **Seaborn** and saw that by Pearson’s method Gross highly correlated with budget.
* Numerized Company names to include it to correlation and see if the second Hypothesis is true. **Unstacked** correlated pairs and show only high correlated ones (more than 0.5), in descending order: Votes and

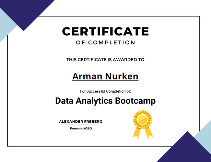
Poetry generator on kazakh language April 2023

* **Data collection** and **Pre-processing**: Poems with avg syllables per line is 11, more than 450 000 characters. Punctuations, specific and irrelevant symbols were removed, all data converted to lower case, did exploratory data analysis by **Word-Cloud**
* Utilized **Tensorflow** to perform **LSTM, GRU** layers and the **model**. Compared two **character-based** architectures. Both have accuracy more than 70%.
* Made survey for the best results, by concluding 62% of surveyors thought that generated poetry is the real one. And by criteria’s like **Rhyme, Rhytm, Meaning** it had points more than average for each

# Education

**Bachelor of Big data analysis** – Astana IT University – Kazakhstan, Astana June 2024

Majors: DBMS, Advanced Python (Deep learning), Statistics and Data Science, Probability and Statistics, Capstone Project (NLP)

**Data analyst bootcamp** – Alex Freberg

Skills: SQL (MySQL), Excel, Tableau, Power BI, Python June 2023