DS & AI job market exploration on hh.ru vacancy data (2020 vs. 2024)

Alexandra Vabnits, Bulat Akhmatov
DS-01

Table of contents

01 Question

02 Motivation

03 Dataset

04 Hypotheses

05 Statistical techniques

06 Results

How have the characteristics of DS and AI vacancy pool changed from 2020 to 2024?

Considering factors such as salary, required experience, and employer regions

Motivation

DS & AI

are highly dynamic, evolving professional fields

Candidates

can adapt their job-seeking strategies effectively

Employers

are able to adjust their hiring approaches and stay competitive

Dataset

Collection

- 2020 dataset collected by Bersenev et al.
- 2024 is collected via hh.ru API request by ourselves.

Preprocessing

1. Unified format.
2. One month cut, only vacancies in Russia.
3. Account 2020-2024 inflation, according to official CPI data in Russia (РОССТАТ).

Search string for API

'Data scientist' or 'Data analyst' or 'ML' or 'AI' or 'Machine Learning' or 'Artificial Intelligence' or 'Аналитик данных' or 'Data Engineer' or 'Инженер данных' or 'Reinforcement learning' or 'Аналитик-исследователь' or 'Нейросеть' or 'Искусственный интеллект' or 'Машинное обучение'

Features of vacancies

salary_from (lower bound)
salary_to (upper bound)
salary_currency
experience_id
area_id
name, etc.

Hypotheses

Comparing 2020 vs. 2024:

Overall salaries

Salaries grouped by required experience

Salaries grouped by areas

Statistical techniques

GoF with KS test

Checking normality of samples.
All samples distribution were not normal.

	Test	KS Statistic	P-Value
0	Salary from 20	0.267606	0.0000
1	Salary from 24	0.185121	0.0000
2	Salary to 20	0.259568	0.0000
3	Salary to 24	0.152430	0.0001

Non-parametric tests for 2020 vs. 2024 comparison

- KS test to check similarity in distributions
- Mann-Whitney U Test with more attention to central tendency

Results | Overall salaries

Mann-Whitney U test, One-tailed

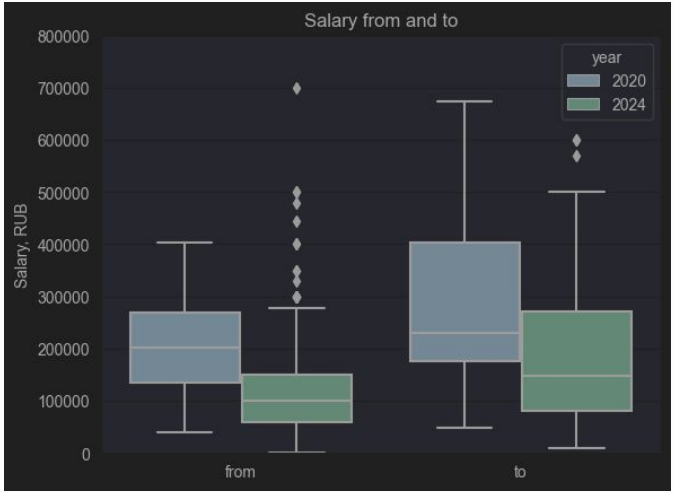
H0: the salaries in 2024 are the same or greater than the salaries in 2020;

H1: the salaries in 2024 are less than the salaries in 2020.

(6 tests total)

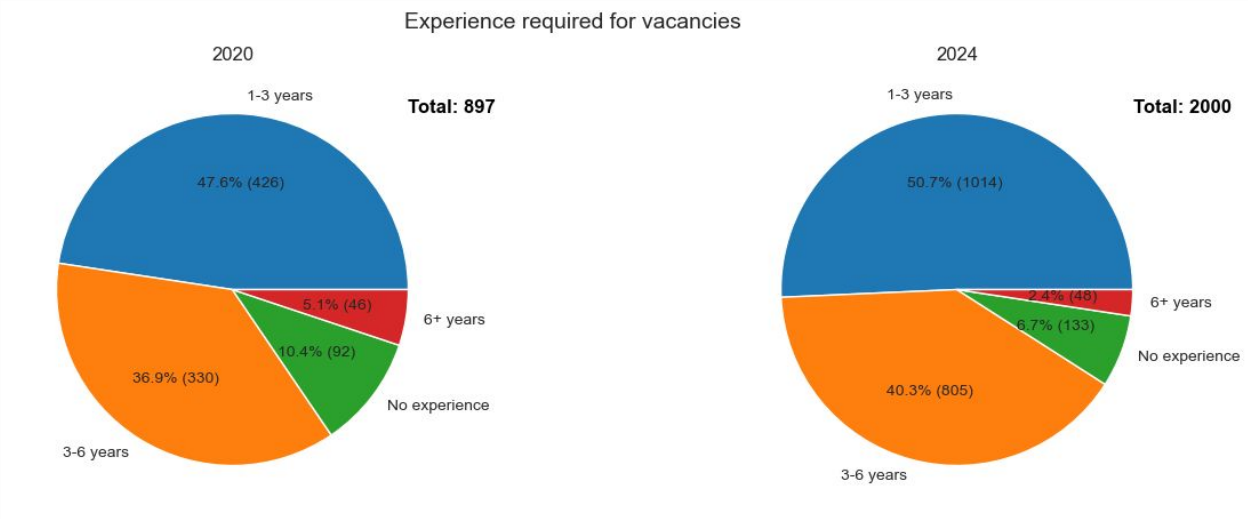
Summary of Statistical Tests

Test	p_value	Stat	Result
Kolmogorov-Smirnov From	9.828036940456279e-18	0.47788857864645656	different
Kolmogorov-Smirnov To	2.6584314675455293e-09	0.40312157721796277	different
Mann-Whitney U From	7.489543387710043e-17	28753.0	different
Mann-Whitney U To	1.0828926210489992e-08	13017.0	different
Mann-Whitney U From One-Sided	3.7447716938550217e-17	8852.0	different
Mann-Whitney U To One-Sided	5.414463105244996e-09	5243.0	different



Conclusion:
the salaries in 2024 are different from the salaries in 2020. Specifically, the salaries in 2024 are less than the salaries in 2020.

Results | Salaries by required experience



2024: less “No experience” and “6+” in favor of “1-3 years” and “3-6 years”.
This change require further research.

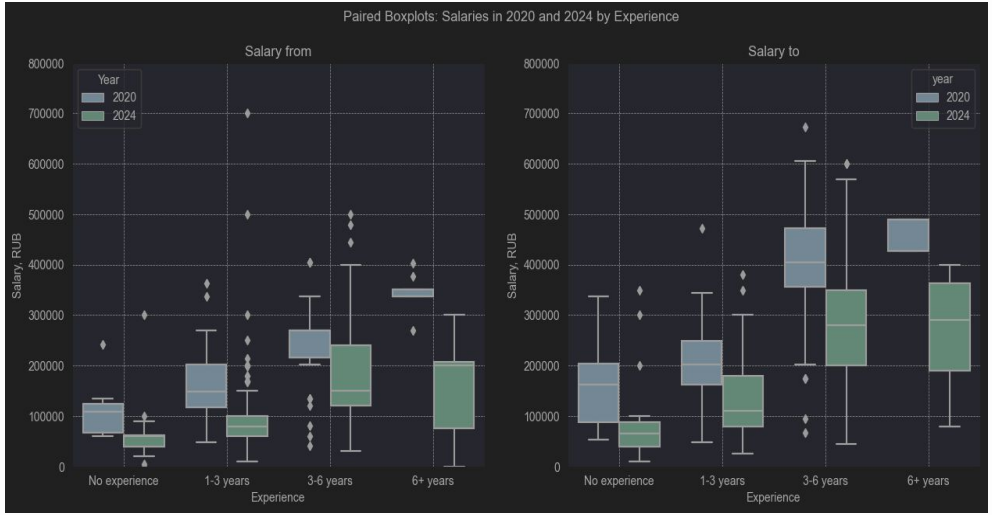
Results | Salaries by required experience

Mann-Whitney U test, One-tailed

H0: the salaries in 2024 are the same or greater than the salaries in 2020;

H1: the salaries in 2024 are less than the salaries in 2020.

Now for pairs grouped by experience (24 tests total)



Conclusion:
the salaries in 2024 are less than the salaries in 2020 for all experience level groups according to one-tailed Mann-Whitney U tests.

Results | Salaries by region

Mann-Whitney U test

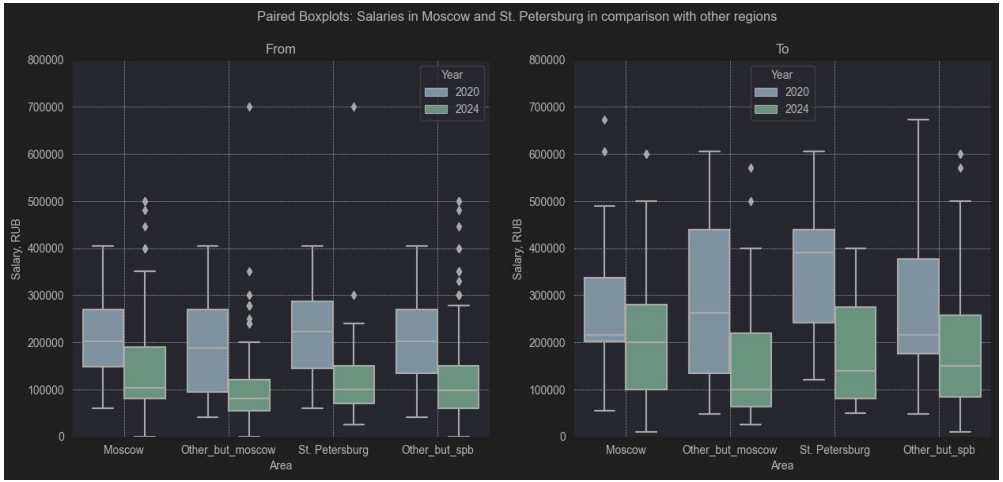
Target cities = {Moscow, St.Petersburg}

Target years = {2020, 2024}

H0: the salaries in *target city* are the same as the salaries in other areas for *target year*;

H1: the salaries in *target city* are different from the salaries in other areas for *target year*.

(4 tests total)



Conclusion:

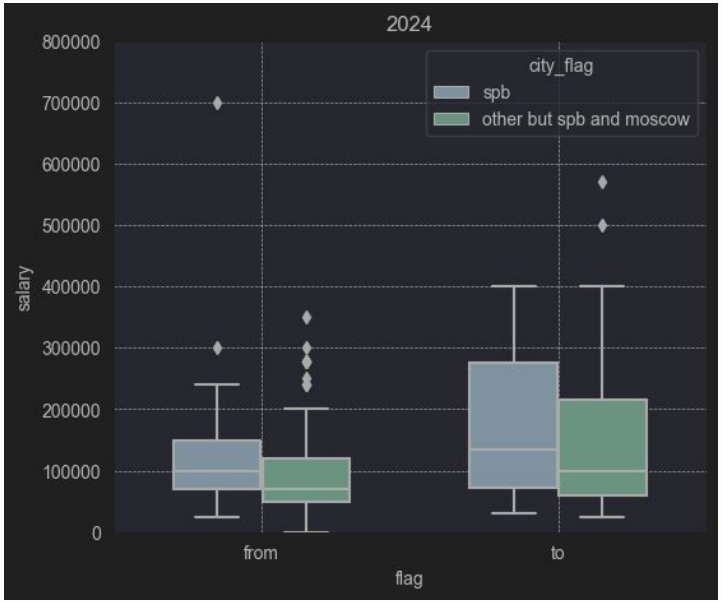
the salaries in Moscow are different from the salaries in other regions for 2024 year. In 2020 for both target cities and in 2024 for St.Petersburg no noticeable difference was detected

Due to lack of data, tests were conducted only for Moscow and St. Petersburg, and only for “salary from” (lower bound for salary)

Results | Salaries by region additional test

Since the number of vacancies from Moscow is much higher than from other regions, the salaries from Moscow have a significant impact on the overall distribution of salaries, so it should be better to compare the salaries in St. Petersburg with the salaries in other regions but Moscow and St. Petersburg

Conclusion:
the salaries in St. Petersburg are different from the salaries in other areas, excluding Moscow and St. Petersburg




```
2024: p-value for St. Petersburg vs. other regions: 0.0027
```

Conclusions

Overall salaries

Salaries in 2024 have decreased compared to 2020 when adjusted for inflation.

By experience

Across all levels of required experience (ranging from no experience to over six years), salaries in 2024 are lower than those in 2020.

By regions

In 2020, there was no significant difference in salaries between Moscow and other regions.
However, by 2024, salaries in Moscow have diverged from those in other regions.

A downward trend in compensation for DS & AI positions is noticeable. Moreover, the difference between capital and other regions has increased.