



# RELEVANCE OF BIG DATA IN BUSINESS

Anna Shakhova

The St. Petersburg State University of Economics,  
Sadovaya street 21, St. Petersburg,  
The Russian Federation

Maxim Naumenko

The St. Petersburg State University of Economics,  
Sadovaya street 21, St. Petersburg,  
The Russian Federation

**Abstract.** The article presents the survey consisting the and author's study of analysis of the scale and structure of Big Data market worldwide within the period of last decade. The authors do perform the self made and forecast of the Big Data market volume and salary in this field for the coming years based on the provided statistics and Prais–Winsten statistics function. Further the authors do perform an attempt to interpret the results obtained taking into account the Figures of the Big Data market development and draw conclusions based on the results of the study. The authors' conclusions boil down to confirming the hypothesis about the growth of the Big Data market, the growth investments as well as growth of salaries level on this sphere within the future, especially for such countries as the People's Republic of India and it's neighbors, as well as the Russian Federation, other countries of the BRICS block. Another main idea is that the salary rate together with the speed of market development are the crucial factors influencing the supplement of the labor force in this segment.

**Keywords:** the Big Data market, labor market, salary rate, the Peoples Republic of India.

## I. THE ORIGIN OF THE PROJECT AND THE TEAM

The team of authors of the article consists of two young daring students, deeply motivated and performing outstanding results. Authors have conducted a survey of the tendencies of the Big Data market Worldwide within the period of the last decade

## II. MATERIALS AND METHODS

The team has conducted the Big Data market and labor market analytics, investigating employment tendencies within the period of the last decade (since 2011 till 2021). The authors team and has implemented the Prais–Winsten statistics function to perform the forecast of the market itself as well as salary rate growth.

## III. METRICS

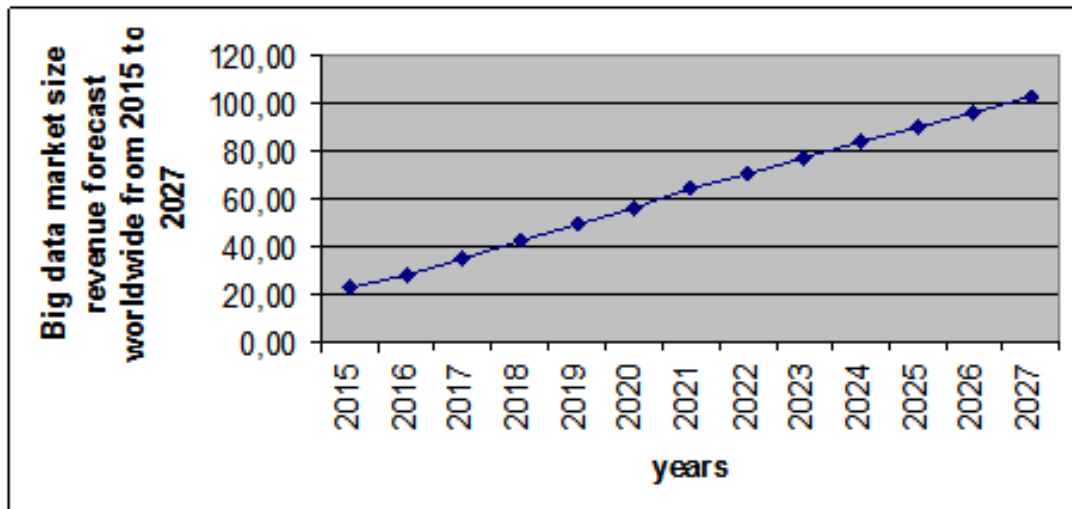
The authors have investigated the time series of the number of employed and the value the salary rate within the period since 2011 till 2021 based on the international statistics data and on given graphs.

## IV. INTRODUCTION

In modern business there are a lot of different trends connected with technologies, particularly using of Big Data in big companies. With development of technologies for data processing appear a lot of opportunities for corporations to use it to increase their profit. So, goal of our research is to analyze market of Big Data in last decade, to investigate how data scientist's influent on business model of their companies and why demand for professions related with Big Data will increase in near future. We have used some econometric models to describe process of growth amount of data scientists and their salary, also we have made forecasts on these trends and in conclusion we say about development of Big Data as one of key parts of modern business and importance for big companies to invest in Big Data technologies and its specialists.

## V. SURVEY

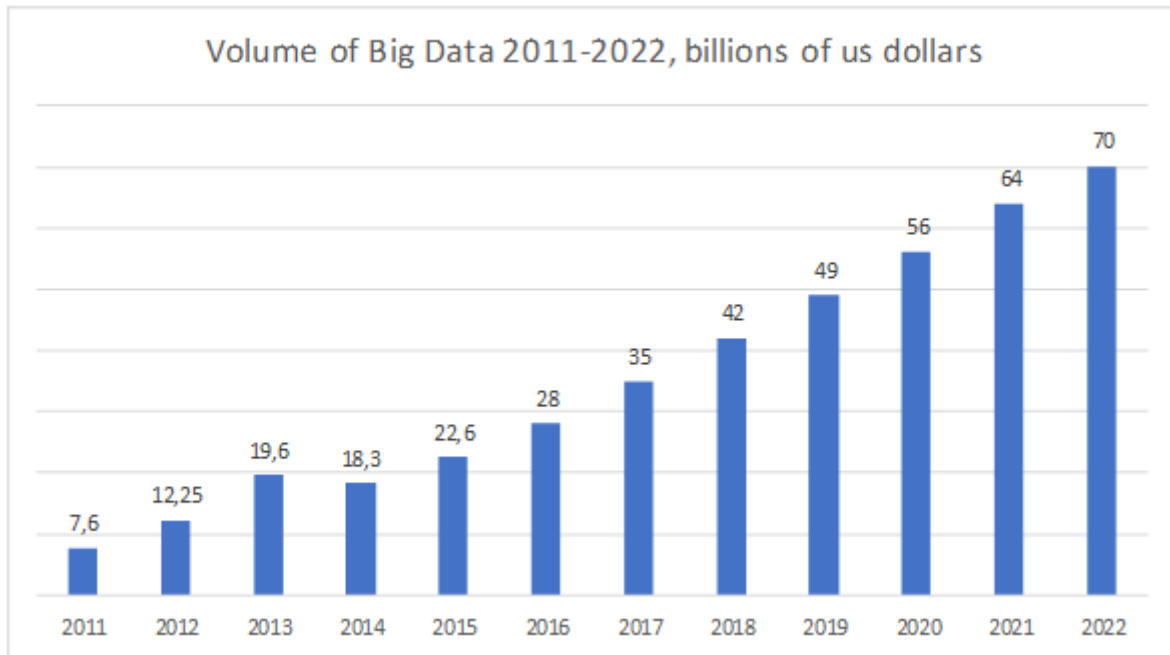
It could be noticed that within the last years attention to Big Data has increased in many times. Why has it happened? What is the reason for it is heard about this industry almost everywhere within last years: in science, business, sport, entertainment, pharmaceuticals retail and other fields? The answer is the following: Big Data it is about information, huge amount of information, which allows to make important decisions. We are now moving actively into the digital world. Information flows in huge streams every second, so there is a need to analyze it in real time to improve the efficiency and security of data.



**Fig. 1.** Big data market size revenue forecast worldwide from 2015 to 2027 (in billion U.S. dollars)

On the Figure 1 here could be seen the tendency of the worldwide Big Data market growth. Traditional analytics is gradually fading into the background, for the reason that it cannot perform so many tasks in a short time. Contrary to traditional database, Big Data is

unstructured and semi-structured information with a weak correlation between the data. If we take a look on market of Big Data in the USA, which is presented at Figure 2, we can notice, that since 2011 volume of this market has surged in 10 times and we can see upward trend.



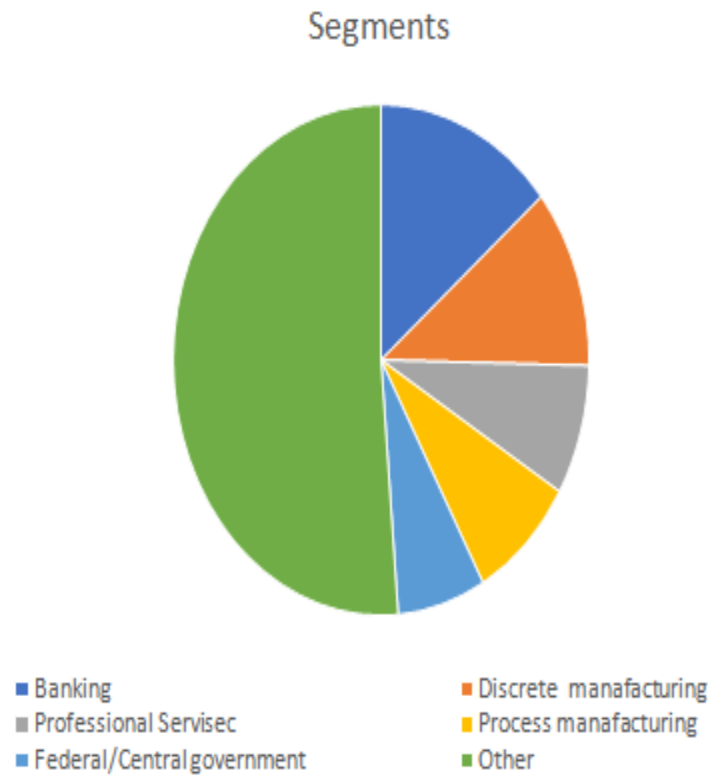
**Fig. 2.** Volume of Big Data 2011-2022 (in billion U.S. dollars) <sup>i</sup>

To describe this trend, we can build econometric model, and if we get a model with interpreted parameters, we will make a forecast for the coming years. We have got a model of Prais–Winsten estimation:  
 $\hat{y} = -11410,4 + 04 + 5,68 * t$   
 t-test (-12,60) (12,64)

The model has parameters:  $R^2 = 0,98$ , F-test = 71,5, DW parameter = 1,62 while  $dL = 0,87$  and  $dU = 1,32$ , normality test = 0,05. These values tell us about a good model and we can make a forecast. According to this model volume of Big Data market in 2023 will be 75,05 billions of us dollars and 80,27 in 2024. It

confirms our assumptions about growth of this market and its relevance. Growth of Big Data volume is predictable, because it has a wide use in many areas and companies Worldwide are in-

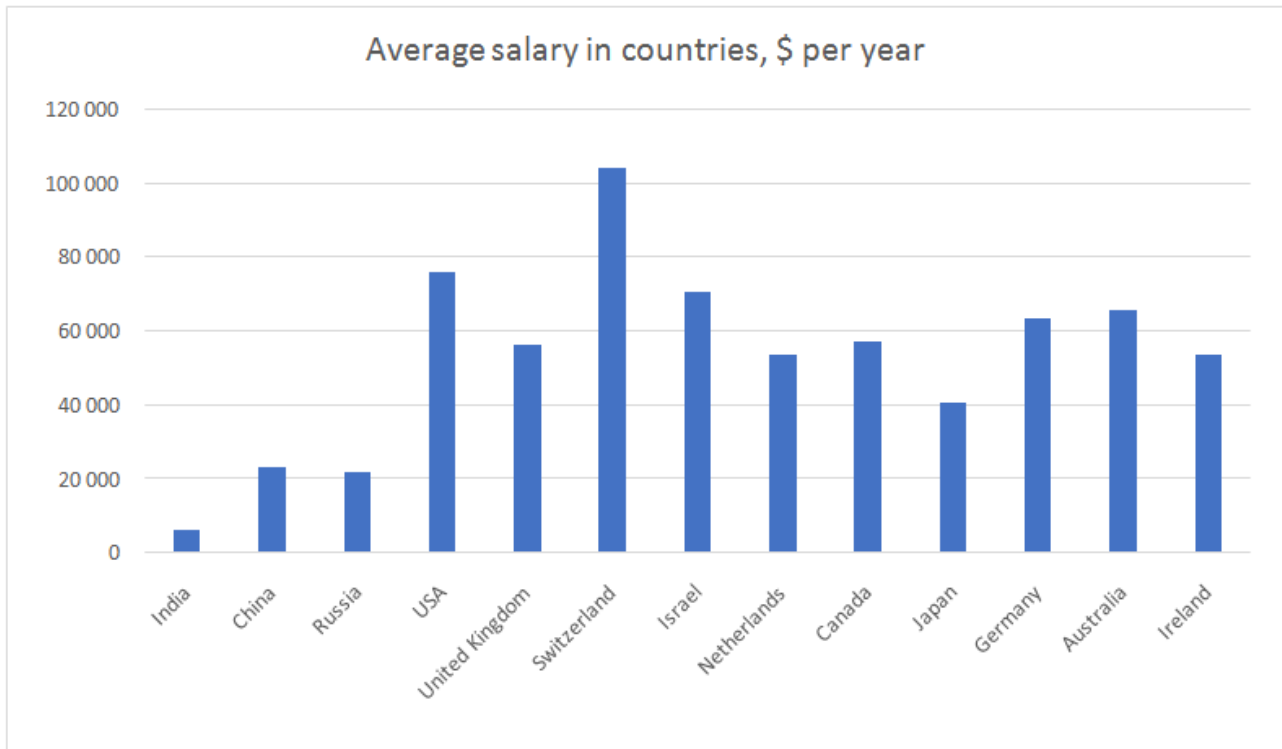
terested to get as much as possible information about us, their consumers. The Figure 3 shows the approximate segmentation of branches where Big Data is used within last time.



**Fig. 3** The approximate segmentation of branches where Big Data is used within last time<sup>ii</sup>

The idea that «the more the companies would know about consumers,- the more money they could earn» works especially in relation to Big Tech companies like Apple, Amazon, Meta, Netflix and Microsoft. Every our step in the Net will be in their services and after its processing, we will be surprising, when algorithms offer us the things we need. Another important fact of active growth of Big Data is, of course, predicting and forecast. Everyone wants to know: what will be in near future, particularly if we talk about top-management of companies. To make decisions, board of directors would ever like to know what results to expect. Science about data is using different instruments and algorithms of machine learning to give a forecast. It is maybe even the key factor of Demand on Big Data.

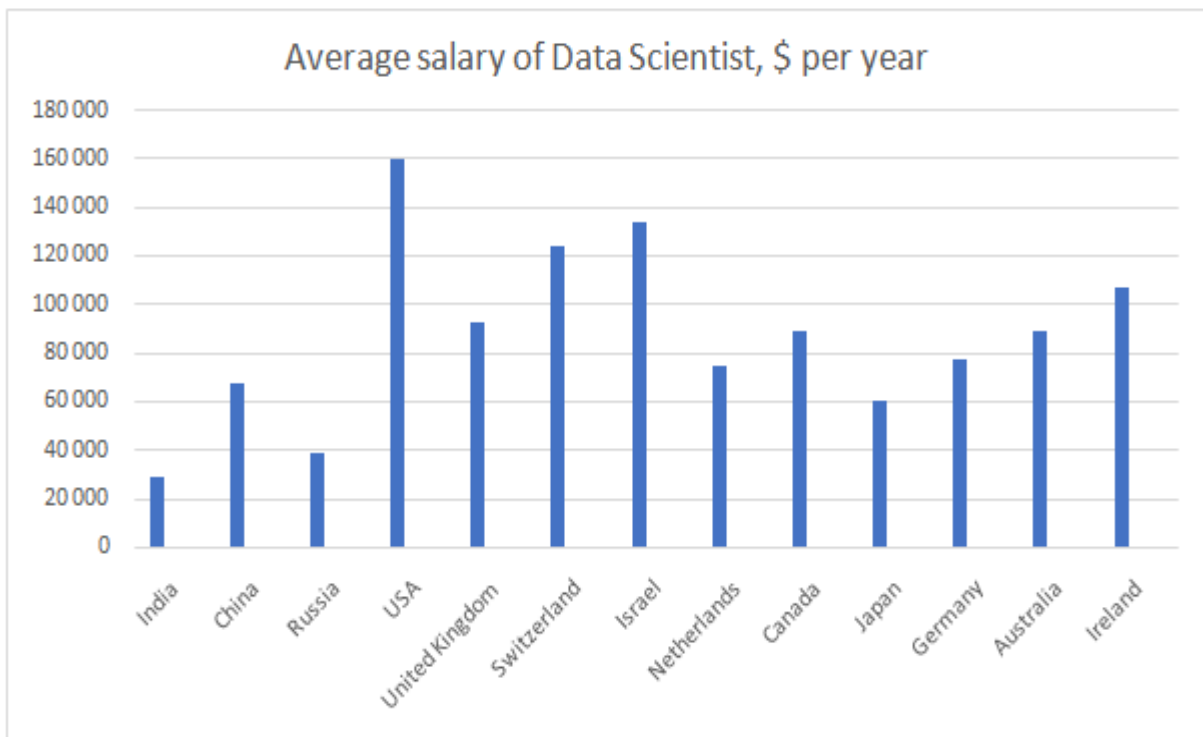
Besides these areas of analyze, different correlations between facts is of no less importance. This area is useful as for business, when there is any correlation, example in consumption and promo action, but the mostly it has significant meaning in science, when scientists need to research influence of one fact on another. Within time there appears more and more connections and data, so, for example, in biotechnologies it is vital part of any research. Due to the need to introduce the use of big data analysis in all areas, there is a high demand for specialists. The labor market in IT specialties has changed a lot over the past decade. To track the dynamics, consider the average wages per year for programmers by country for 2014.



**Fig. 4.** Average salary in countries in 2014 (the US dollars per year)<sup>iii</sup>

In relation to the average salary in the country, India, China and Israel were the leaders. And this is not surprising, because India in the early 2000s became the leader in terms of

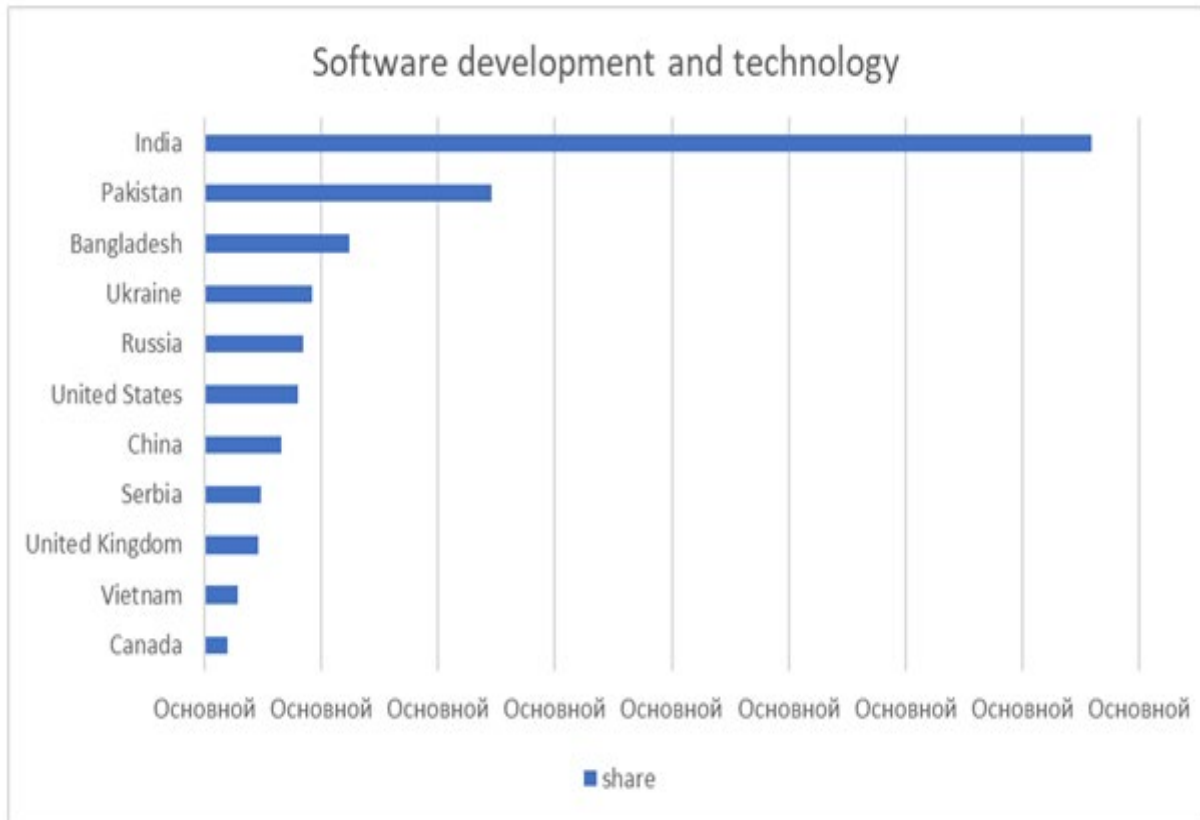
IT technology exports and occupied a dominant position in the IT outsourcing market after the Y2K problem.



**Fig. 5.** Average salary of Data Scientist in countries in 2014 (the US dollars per year)<sup>iv</sup>

Every year India invested more and developed in this direction, there were a lot of foreign orders, and the workforce too. The number of orders increased, many companies used the services of Indian programmers, because it was quite cheap relative to Western standards. The cheapness of Inter-

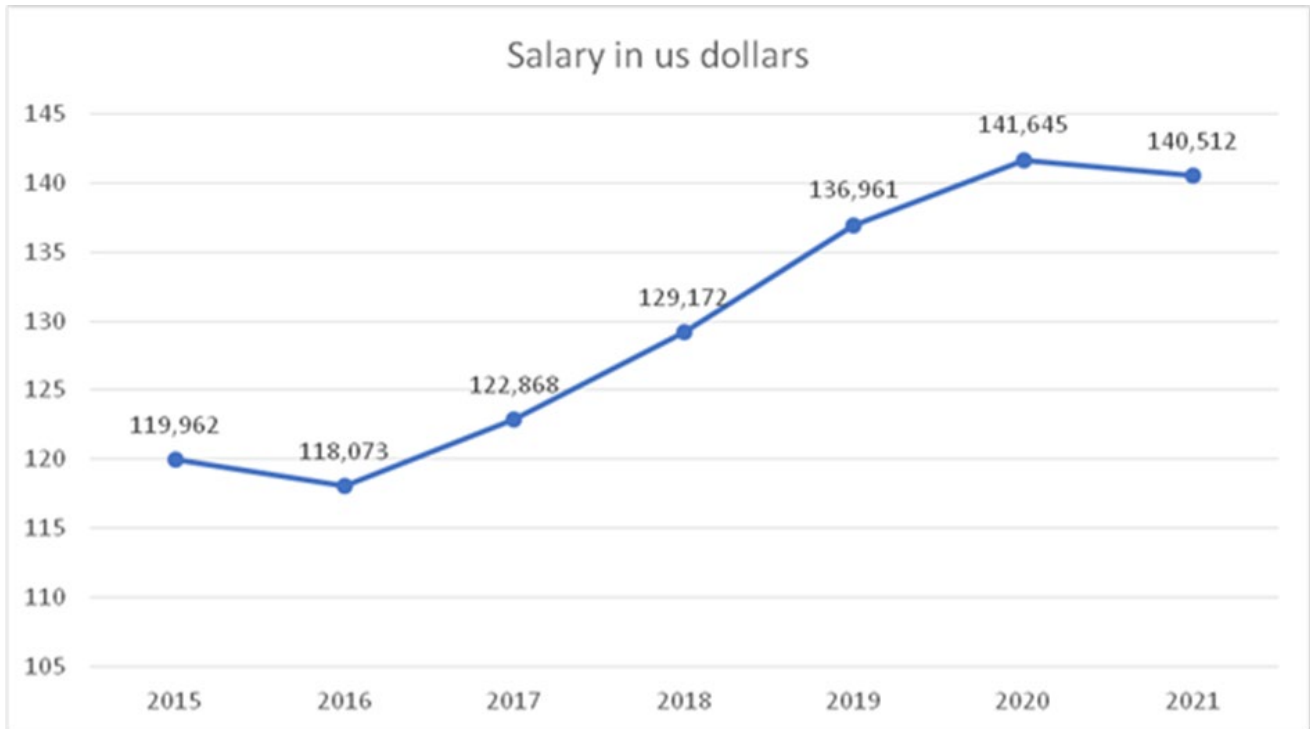
net access and locally made smartphones also had a big impact. The consequence of all these factors is that India has strengthened its position in the IT outsourcing market, which is illustrated by the Figure 6.



**Fig. 6.** Software development and technology according to Online Labor Index chart<sup>v</sup>

The chart is created by the Online Labor Index developed by the International Labor Organization and the Oxford Internet Institute of the University of Oxford. It shows the offer of online freelancers in different countries and professions, tracking the number of projects and tasks on different platforms in real time. The data covers the period from 2017 to 2023. The leaders are India and its neighbors in South Asia - Pakistan and Bangladesh. The ranking includes

Ukraine, Russia, USA, China, Serbia, United Kingdom, Vietnam and Canada in descending order. Over the years, the situation on the labor market in the field of IT technologies has changed for the better in terms of market and salary growth. At the moment, the need for specialists is accompanied by a fairly high salary level, which is presented further on the Figure 7.



**Fig. 7.** Salary in IT technologies sector: the dynamic since 2015 to 2021<sup>vi</sup>

This suggests that companies are willing to pay a high price for increasing the efficiency of the process. The leaders in remuneration were USA, Israel and Switzerland. If we look at Data Scientist's salary over time, we will see, that their salary grows, increased from 2014 to 2021 year on 20 thousand dollars.

We also can build econometric model, to describe this process and to predict situation in next years:

$$\hat{y} = -8726,76 + 4,39 * t$$

t-test (-7,257) (7,365)

The model has next parameters:  $R^2=0,91$ ,  $F=54,25$ ,  $DW=1,71$  while  $dU$  and  $dL$  do equal 0,7 and 1,35 so we can say, that this model can be used to describe situation on labor market and we have opportunity to make forecast. According this model, in 2023 average salary on data scientist market will be 151,829\$.

If we are researching the labor market of a Big Data analyst, then it is worth mentioning the main salary criteria for all countries. In the field of programming, huge differences in the income of specialists make up categories that characterize the amount of knowledge and experience: Junior, Middle, Senior. Work experience has the greatest impact. Often a Junior is a programmer with up to a year of experience while they are on an internship or under the guidance of a more experienced specialist. Middle - from 1 to 3 years old, they can perform more complex tasks on their own. Senior is an experienced developer who manages projects and teams. The location of the employer affects in this way: the closer the job is to megacities and large cities, the higher is

the salary. The programming language that the developer owns also has an impact. Based on vacancies on websites, we can conclude that the most in-demand skills for Big Data for an employer at the moment are: python, sql, git, followed by big data, deep learning and classic machine learning. Next - some more highly specialized tools, devops (development and operations) and development. An additional advantage in favor of higher incomes is also the desire and willingness of a specialist to cope with various tasks. The ability to be multifunctional and adaptive will be a good advantage. And, of course, it is necessary to take into account that in large companies the salary is always higher, but it is more difficult to get a job there, since the larger the scale of the company is, the more requirements for the developer do exist.

## VI. RESULTS

Upon the investigation there several results could be concluded.

The first one is that all in all, situation in Big Data sector is predisposing for investments and career. A lot of companies use data to optimize business-processes and increase the cost. Within the last time we can notice interest of not only tech companies like Google or Microsoft, however also of the banking sector, which starts investments in this field. So Big Data scientists soon will be demanded in every sphere. Every big company is interested in development of these technologies and they are ready to hire more and more perspective data scientists from all over the world. Salaries and



demand of data scientists are growing in fast temps and in near future this trend will be continued, which makes choice of career easier, if one wants to be in demand worker.

Another statement is that the further growth of the Big Data market should become the very important factor of development of economics of the countries of the BRICS block, including the People’s Republic of India, first of all. To our mind the same statement is fair for the Russian Federation as well.

#### VII. REFERENCES

1. Oesterreich, Thuy Duong, Eduard Anton, and Frank Teuteberg. “What translates big data into business value? A meta- analysis of the impacts of business analytics on firm performance”. *Information and Management* 59.6 (2022): 103685
2. Ajah, Lfeyinwa Angela, and Henry Friday Nweke. “Big data and business analytics: Trends, Platforms, success factors and applications”. *Big Data and Cognitive Computing* 3.2 (2019): 32.
3. Quan, Tee Zhen, and Mafas Raheem. “Salary Prediction in Data Science Field Using Specialized Skills and Job Benefits- A Literature”. *Journal of Applied Technology and Innovation*( e-ISSN: 2600-7304) 6.3 (2022):70.
4. King, John and Roger Magoulas. 2015 data science salary survey. O’Reilly Media, Incorporated, 2015.

---

<sup>i</sup>The same source

<sup>ii</sup>Ajah, Lfeyinwa Angela, and Henry Friday Nweke. “Big data and business analytics: Trends, Platforms, success factors and applications”. *Big Data and Cognitive Computing* 3.2 (2019): 32.

<sup>iii</sup>Quan, Tee Zhen, and Mafas Raheem. “Salary Prediction in Data Science Field Using Specialized Skills and Job Benefits- A Literature”. *Journal of Applied Technology and Innovation*( e-ISSN: 2600-7304) 6.3 (2022):70.

<sup>iv</sup>The same source

<sup>v</sup>The same source

<sup>vi</sup>The same source