**How to Enter Data Era?**



# Table of Content:

1. Introduction
2. What's the Best Tools and Knowledge to enter the Field?
3. Let's predict the Best Job title based on your Skills.
4. What Affect the Salary?
5. Conclusion.

## **Introduction**

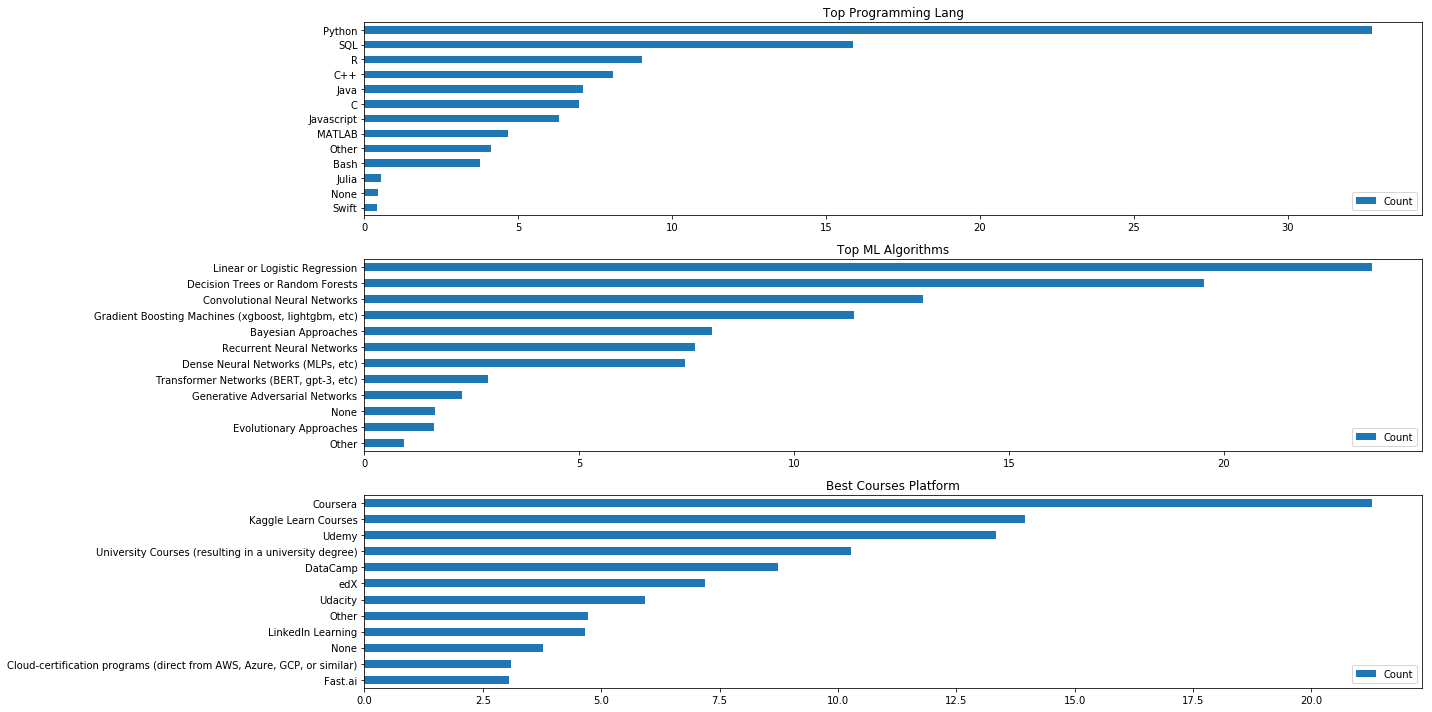
Are you asking yourself about what’s needed to Enter Data Era? While you stuck about what are different titles available and what’s different between each of them.

Also, asking about Needed Skills to enter the Field and What will be trend for the Next 2 Years. And How to choose your salary Range widely.

Now, Let’s listen from Experts in the field with brief analysis for [Kaggle Survey 2020](https://www.kaggle.com/c/kaggle-survey-2020/overview) for +20K respondent from all over the world.

## **What's the Best Tools and Knowledge to enter the Field?**

To answer this question, we need first to know where we will find best Theoretical background and Courses.

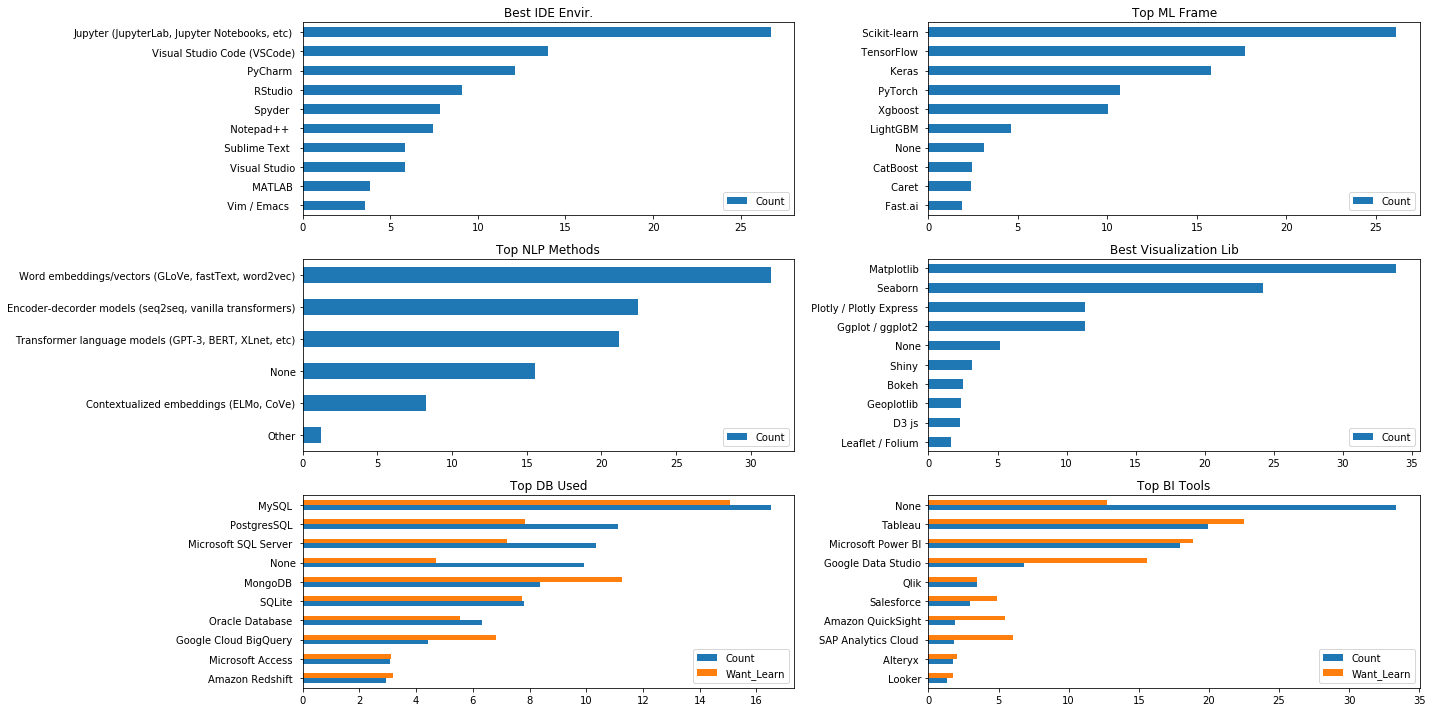


We’ll find that more than **quarter** of respondents nominate “**Coursera**” to take their courses on then “**Kaggle** Courses” and “**Udemy**”. These three platforms are almost best Platforms for +10K people around the world.

Now, Let’s See what to target inside these platforms.

Hot Programming Languages to target are **Python**, **SQL** and **R** with **35%**, **17%** and **9%** correspondingly.

The most popular Machine Learning Algorithms are “Linear/Logistic Regression” which used at our analysis later with **25%** then “Random Forests” and “Convolutional Neural Networks”.



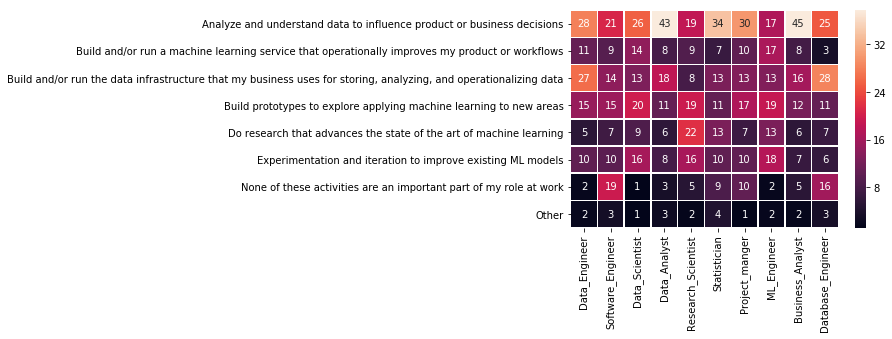
So, What’s the Best IDE Environment to test our code on? **25%** recommend **Jupyter** Platforms and labs to work on then Visual Studio and **PyCharm**.

After all of that we need to present and Visualize our finding, we will find that **MatplotLib** and **Seaborn** take the most recommendation with **+55%.**

With the more complex of our Data and huge need for Visualization specific tools to work with, we will find that **currently** almost **35%** of Users doesn’t need any Extra Tools but almost **25%** of them will give a try for **Tableau** Visualization Tool in the Near future.

## **Let's predict the Best Job title based on your Skills**

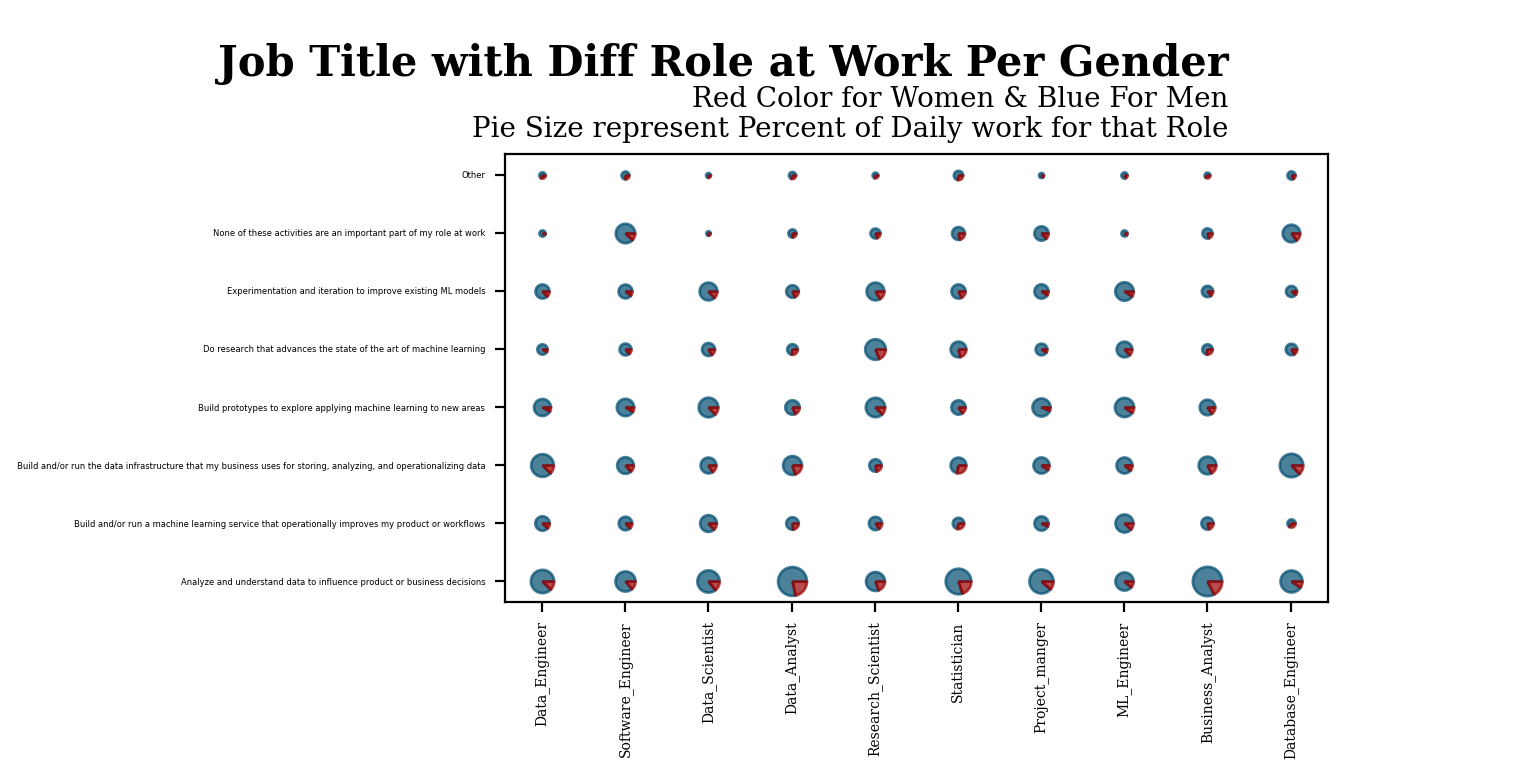
After Successfully identifying Needed Skills and hot Tools and Algorithms, Let’s go deep inside different Job Titles and Know what rules you will do.



From above Heat Map, we can easily identify that we have 10 Job Titles deal with Data. Below some Insights about each of them:

* Data Engineer: **+50%** of his time will analyze and understand data to build use cases.
* SW Engineer: Will analyze and understand data to do other jobs not included at our Survey but could be included in Building SW packages will work as base for Other jobs.
* Data Scientist: We will see that it’s time scattered between almost all of rules, This’s because he is the last step at job hierarchy which get all deep insights from data.
* Research Scientist: **+40%** of his time doing advanced searches to improve Current algorithms and exploring New areas of Interest.

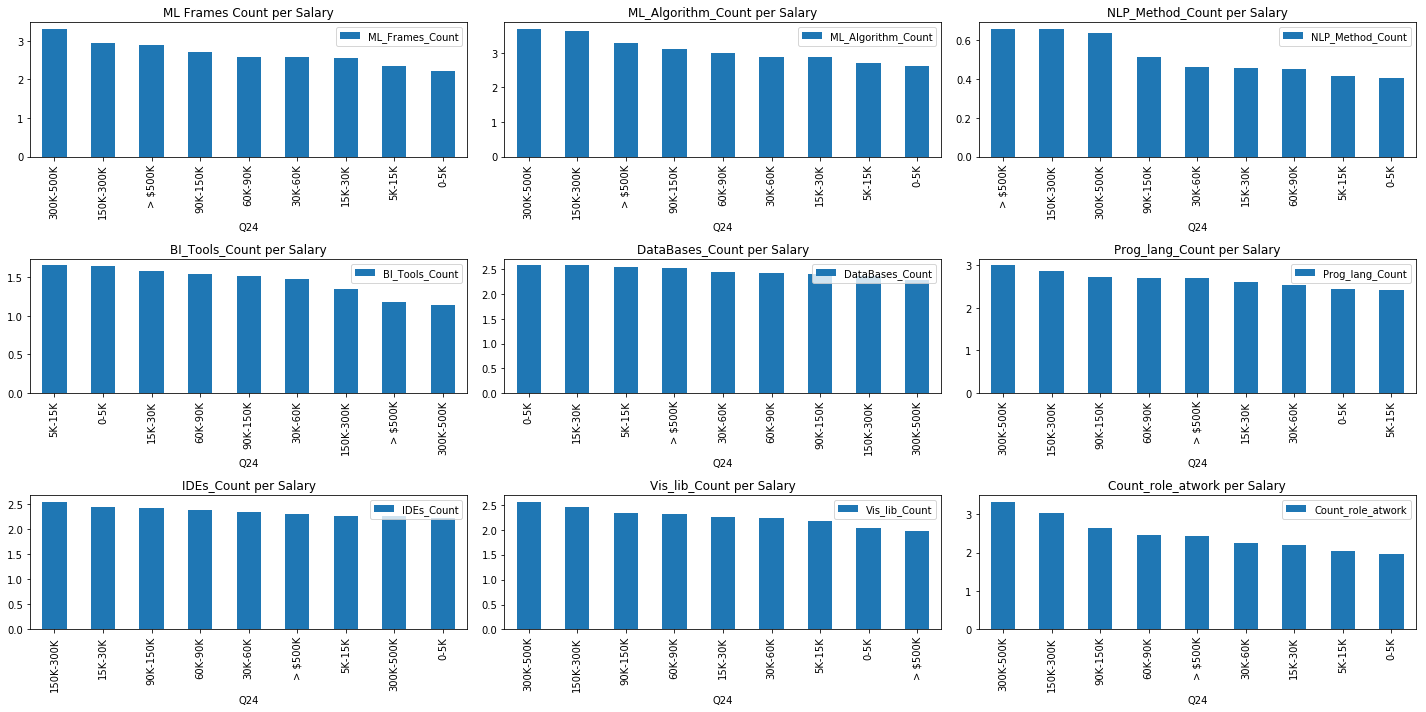
Now, Let’s See the Power of “**Women**” inside each of these jobs.



Above 3D Map, represent Work load of each Job Title with different Job Rules “Like a/m heat map” but with including percent of Women “**Red Colored** inside Pie”. Get your own Insight from it ☺

## **What Affect the Salary?**

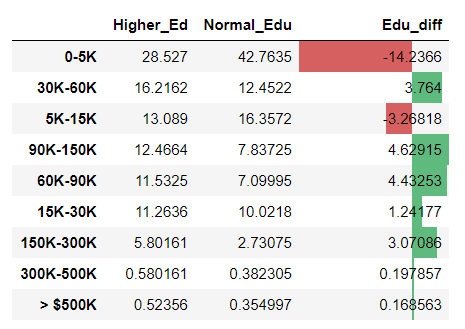
Let’s correlate Salary ranges with the tools and Skills you learnt from First Section.



From above charts we can see the below strong Correlations:

* The More ML Algorithms, Frames and NLP Methods the More Salary you will be paid.
* However, the importance of Business Intelligence and DBs, It’s not directly reflect your annual Salary.
* The “One To One” relation of count of rules you do is still exist and affect your Salary till 500K $ annually, After that range All of Us need to know what to do to gain more than It ☺

The hottest question all the times and never Expired, Is the Higher Formal Education I get The more I paid??



From the above chart, we will find that **+80%** of Normal Education paid less than **60K $** annually on other hand only **60%** of Higher Education paid at same Salary range.

## **Conclusion**

Let’s Conclude the Steps you need to follow to enter the field with best Salary range for you:

1. Go to one of hot Courses platforms to take important knowledge.
2. Let’s target **Python**, **Linear**/**Logistic** Regression as a start to the field.
3. Practice more on **Jupyter** IDE platform and use one of these Visualization libraries **MatplotLib** and **Seaborn.**
4. Based on your preferences target one of Job titles and see what’s the job rules you will do from Section 2 at this Post.
5. After that, you need to Know that “**Studying Never Die at Data Era ☹**” you need to advance your career with more algorithms and methods or targeting Higher Formal Education to be paid more.

Hope that study and this analysis are helpful and interesting for you.

This analysis done under [Data Science Nanodegree from Udacity](https://www.udacity.com/course/data-scientist-nanodegree--nd025).

You can access the full code used at this study [here](https://github.com/KhaledGamal1991/How-to-Enter-Data-Era).