

BUSINESS INTELLIGENCE – PROJECT (Power BI & Tableau)

ANALYSIS ON TEACHERS IN PAKISTANI UNIVERSITIES

Name: Syeda Faakeha Ahmed

ERP: 18603

Dataset: Pakistan Intellectual Capital (*Downloaded from Kaggle*)

As cited by this collaborator on Kaggle,

“According to HEC, Pakistani universities are producing over half a million graduates per year, which include over more than 10,000 Computer Science/IT graduates.”

This lays out an important question for us regarding the teachers who are training these students for their better future. We know that CS/IT, which is a vast field, is undoubtedly gaining popularity due to the increase in technological advancements in this world. Hence, the number of Computer Science graduates are alarmingly increasing, not just in Pakistan, but all over the world.

I took this opportunity to find out more about the teachers and professors who are the masterminds behind the success of these CS graduates. This dataset dates back to 2014-2015.

INTRODUCTION

This dataset introduces information about teachers across multiple universities in Pakistan, both in the public and private sector. Their names, qualifications, current status, interests and expertise etc. This will be discussed in detail in the sections below.

PROBLEM STATEMENTS

1. To understand the dynamics of CS/IT teachers in Pakistan such as their educational information and their current teaching status. To observe patterns between the teacher's graduated field and their current teaching area/field.
2. To explore information about the teachers and universities across different provinces.

DATA TRANSFORMATION (Power BI)

○ COLUMNS REMOVED / DELETED

I have removed the following columns from my dataset since the data in them is messy and unstructured and a separate analysis needs to be done on them through text analysis.

- ‘Area of specialization/ Research Interests’
- ‘Other interests’

Multiple columns can be derived from these and those will have to be analyzed in a detailed manner to structure the data.

I have also removed 'Teacher Name' because it cannot be analyzed. However, before deletion, I have created a new column 'Gender' from it which will be discussed below. I have also removed 'Year' since it contained many null values and it did not fit into my thought process for analyzing.

For now, I will focus on the teachers' qualifications and current teaching status.

- **MISSING VALUE ANALYSIS** (As mentioned in Data.Info)

'Teacher name', 'University Currently Teaching', 'Department', 'Province University Located' and 'Designation' have no null values present.

Since this is a purely dimensional-based dataset, I could not remove or delete any rows with null values. Whatever data is present had to be worked with for a complete analysis. Having this in mind, I replaced the null values of columns 'Terminal Degree', 'Graduated from' and 'Country' with "Did not specify" to acknowledge that for any reason, the teacher did not want to share this information. It is also possible that **HEC (Higher Education Commission of Pakistan)** could not find information regarding these fields since they are the ones who have accumulated this data.

- **CONSISTENT CATEGORICAL VALUES**

It can be observed that the data is consistent enough, throughout the entire dataset. This might be the case since it was accumulated and is maintained by **HEC (Higher Education Commission of Pakistan)**. However, I did tweak the data a little according to what I thought would give me a better analysis.

- In the 'Department' column, there was repetition in the way that the departments were stated. I replaced these values with 'CS/IT' since this looks cleaner and they all have the same meaning.
 - "Computer Science and IT",
 - "Computer Science & IT",
 - "Computer Science and Information Technology"
 - "CS&IT"

Furthermore, I also replaced these values with 'CS' to make it shorter and easier to read.

- "Computer Sciences"
- "Computer Science"

For the same reason mentioned above, I replaced these values with 'IT'.

- "Information Technology"
- "School of Information Technology"

○ **ADDITIONAL COLUMN**

- In the '*Terminal Degree*' column, the values were repetitive and inconsistent, so I created a new column named '*Basic Degree*', that holds values 'Bachelors', 'Masters', 'PhD' and 'Post Doctorate' according to the values given in the '*Terminal Degree*' column.
- Created a new column '*Gender*' (*mentioned in the 'Columns Removed/Deleted' section above*) which contains the values:
'Male' where the '*Teacher Name*' column contains "Mr.",
'Female' where it contains "Mrs." or "Ms." and
'Cannot be determined' where the '*Teacher Name*' contains anything other than the values above such as "Dr.", "Engr.", or just the actual name of the teacher since it would be difficult to determine the gender based on these values. This will provide us with a general insight into the gender ratio of CS teachers in Pakistan.

DIMENSIONS:

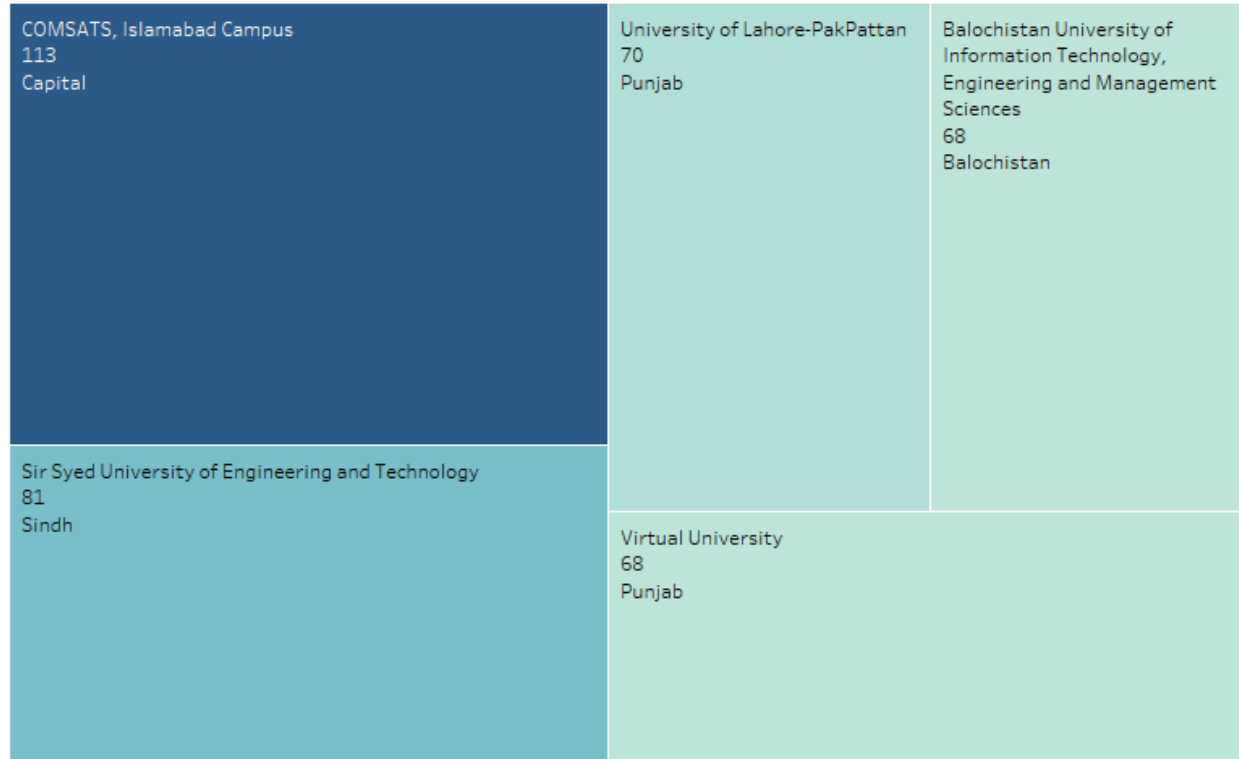
I am left with the following columns for analysis:

- University Currently Teaching
- Department
- Province University Located
- Designation
- Basic Degree
- Terminal Degree
- Graduated from
- Country
- Gender

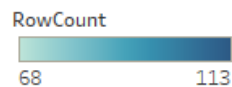
ANALYSIS (Tableau)

Following the queries mentioned in *Data.info*, the charts produced are shown below.

top 5 universities with count



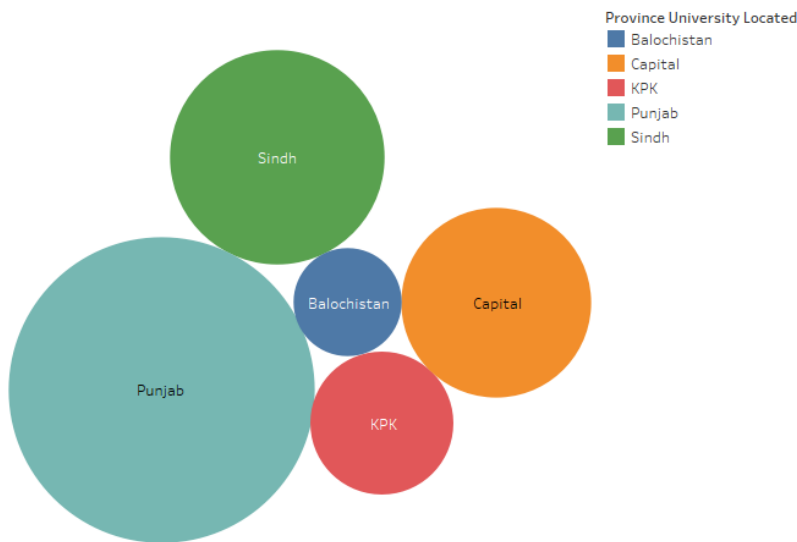
University Currently Teaching, RowCount and Province University Located. Color shows RowCount. Size shows RowCount. The marks are labeled by University Currently Teaching, RowCount and Province University Located. The view is filtered on University Currently Teaching, which has multiple members selected.



This is the view for the top 5 universities in Pakistan based on the count of CS teachers. COMSATS, Islamabad campus has the highest count of 113 teachers and is leading by a relatively large margin.

We can see that there is at least one university from each province present in the top 5 except for KPK. This indicates a shortage of CS teachers in this province. Punjab has 2 major universities with a large count of CS teachers, so it is possible for Punjab to have the highest count overall. But we will analyze this below.

province wise count



Province University Located. Color shows details about Province University Located. Size shows RowCount. The marks are labeled by Province University Located.

We can see from the graph that Punjab has the highest count of teachers compared to the other provinces. Balochistan has the least count of teachers. Although we assumed a shortage of teachers in KPK, it can be seen that the situation is worse in Balochistan.

According to a resource (mention in the sources section), *“Balochistan's government has always been neglecting the decades-long education crisis in the province, and has never been sincere in efforts to increase enrollment of out of school children.”*

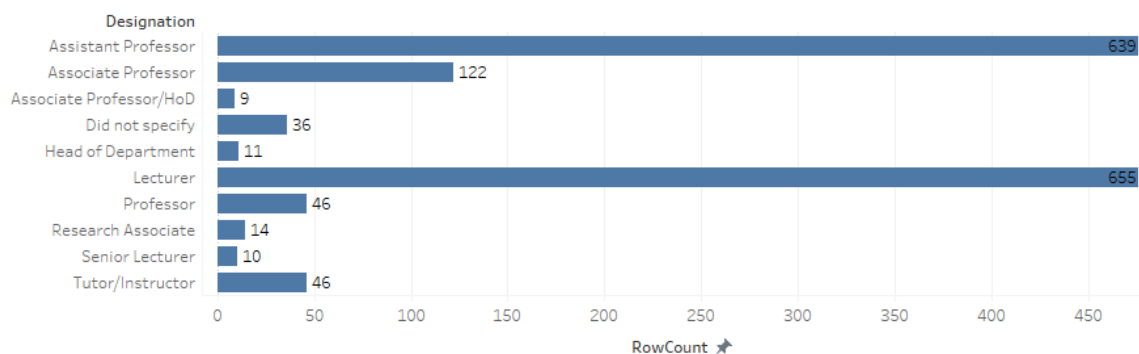
department wise count

Department	
CS	1,152
CS/IT	150
Computer Engineering	81
IT	81
Computing	41
Computing & Information Sciences	38
Institute of Mathematics & CS	29
CS and Engineering	25
CS & Engineering	23
FACULTY OF COMPUTING & ENGINEERING	22
Computer and Information Sciences	20
Department of Computer Systems & Software Engineering	16
CS and Software Engineering	14
ENGINEERING, SCIENCE & TECHNOLOGY	14
Software Engineering	1

It can be observed that most universities have the department name as CS or Computer Science but there are many variations including Computer Engineering, CS/IT, Information Sciences etc.

There are only 3 universities whose pure focus is on “Software Engineering”. These are “Fatima Jinnah Women University”, “International Islamic University, Islamabad” and “Mehran University of Engineering and Technology, Jamshoro”.

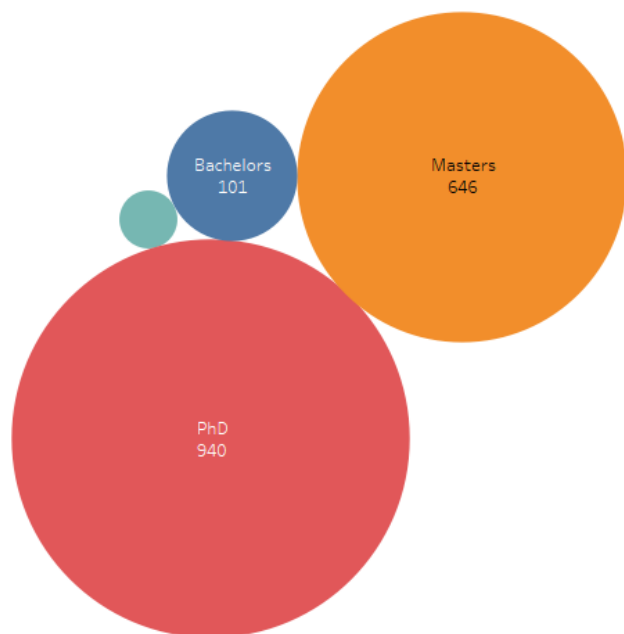
designation wise count



This chart shows the top 10 designations found in the dataset.

From this chart, we can see that more than 1000 teachers are either “Assistant Professors” or “Lecturers” which makes sense based on what the data is about. However, there are some other designations such as “Head of Department” with a count of 11, “Research Associate” with a count of 14 and so on. 36 teachers did not specify their designation.

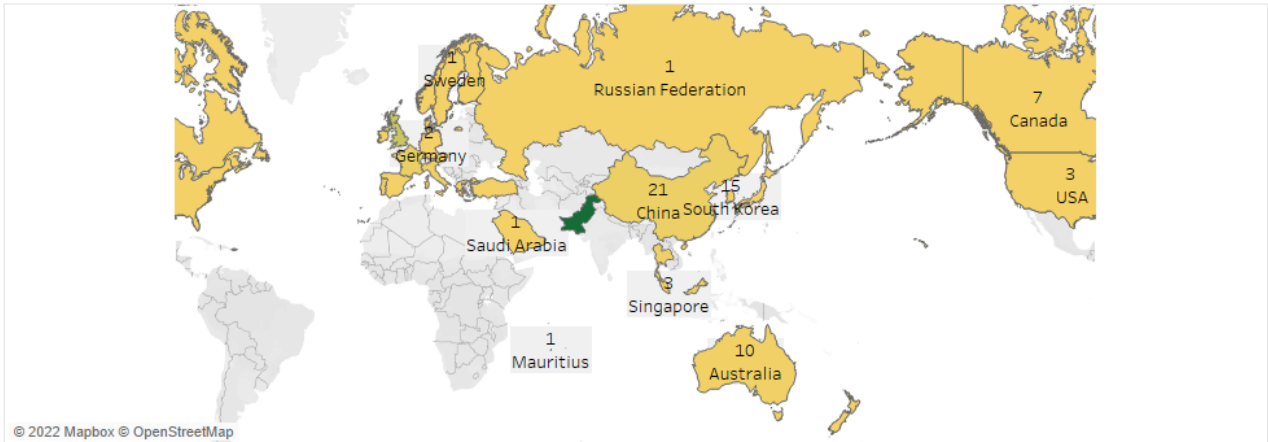
degree wise count



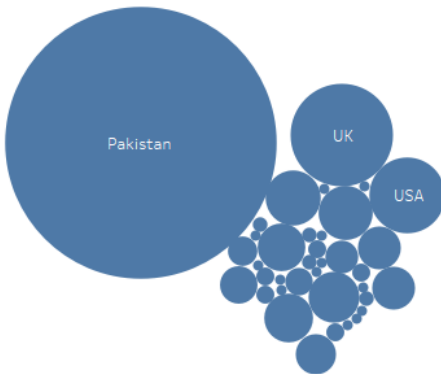
This chart shows the number of teachers holding either a “Bachelors”, “Masters”, PhD”, or “Post Doctorate” degree. I have used the newly created column ‘*Basic Degree*’ for this purpose to easily distinguish between the degrees. We are easily able to tell which degree stands out the most for university professors.

Comparison between ‘*Designation*’ and ‘*Terminal Degree*’ is discussed in detail later in this report.

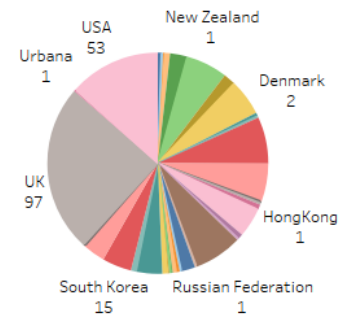
country graduated count



country graduated count2



country graduated count4



These charts give a rough estimate of the teachers that have graduated from Pakistan vs those who have graduated from universities abroad. The map shows that these teachers have graduated from many different countries of the world, even though the count of teachers is not much in such countries. 681 teachers have graduated from Pakistan as can be seen above, which is leading with a huge margin as compared to the other countries.

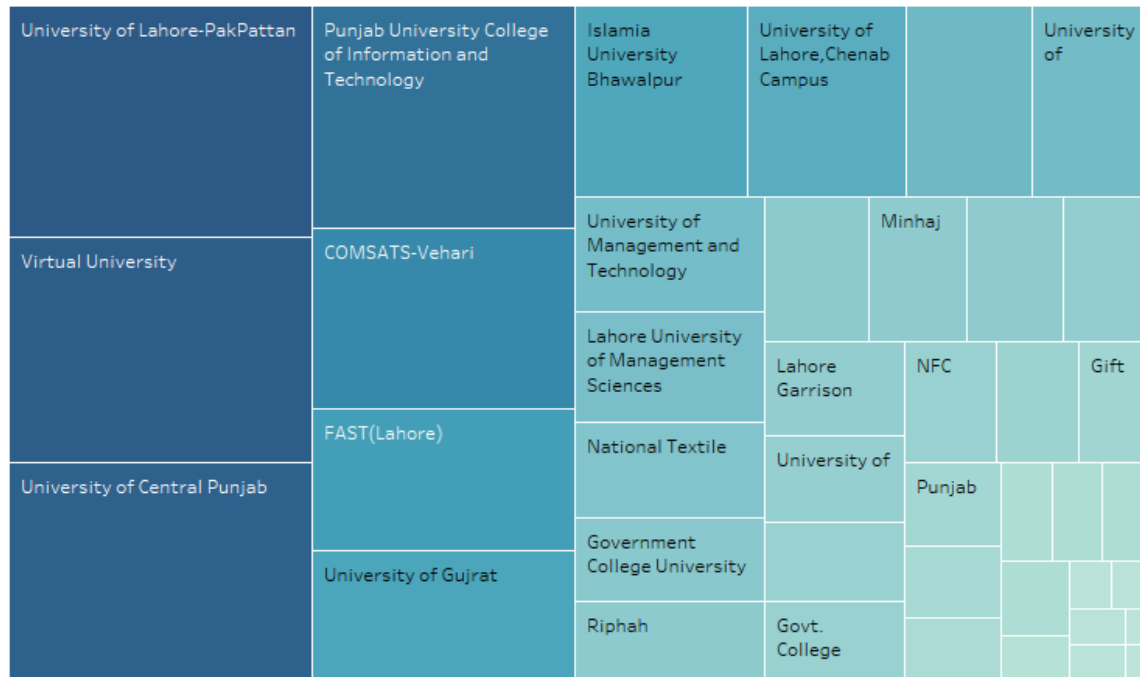
So, I removed "Pakistan" by filtering it out to see what the leading countries are. I also filtered out "Did not specify". We can see the count of other leading countries in the pie chart given in the above dashboard. UK is leading in this category, followed by USA, Germany, Thailand, Spain, South Korea and so on.

FINAL LIST OF BI QUERIES

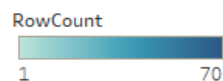
- 1. How many distinct universities are present in each province? Which province has the highest count of teachers? Which province is lacking?

As we discussed above, Punjab has been found to have the highest count of teachers in Pakistan and this can be shown below with the number of universities only located in Punjab.

punjab teacher count

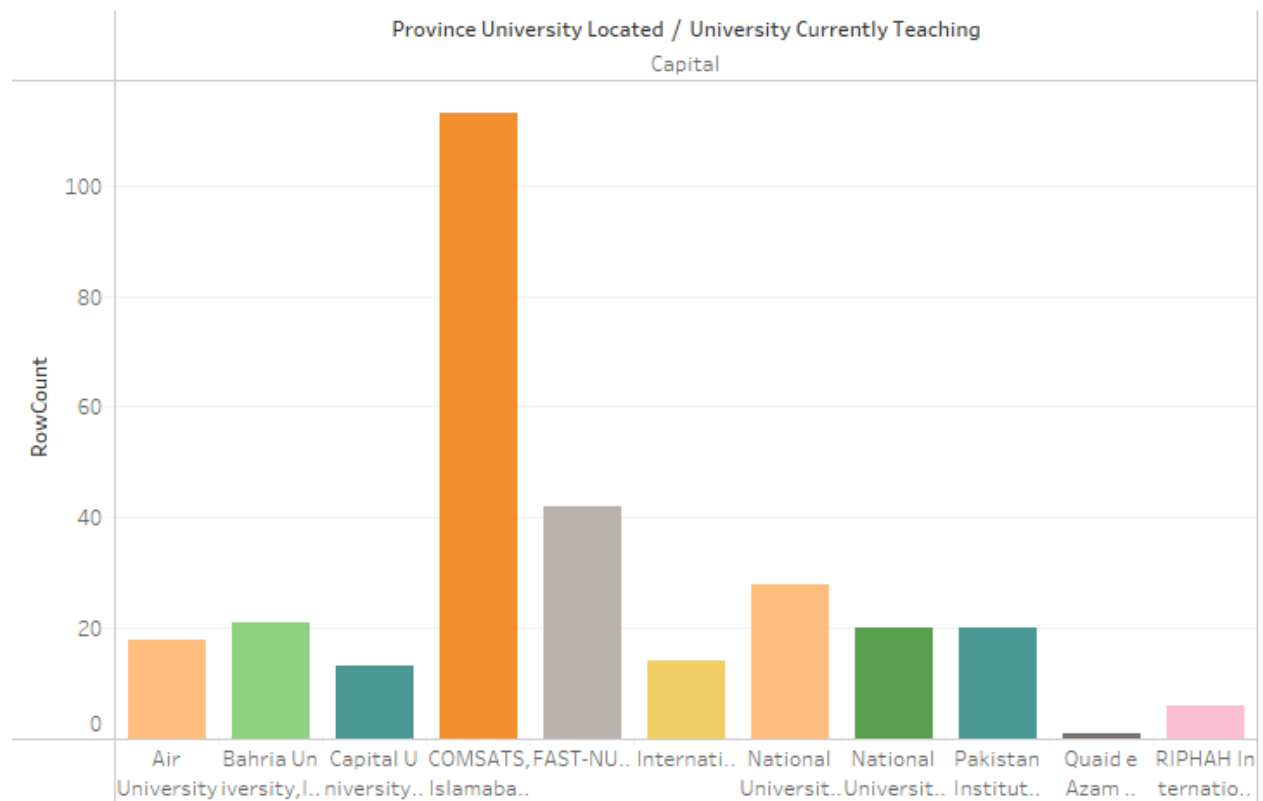


University Currently Teaching. Color shows RowCount. Size shows RowCount. The marks are labeled by University Currently Teaching. The data is filtered on Province University Located, which keeps Punjab.



We can see that University of Lahore-PakPattan, Virtual University, University of Central Punjab have the greatest number of CS teachers.

capital teacher count



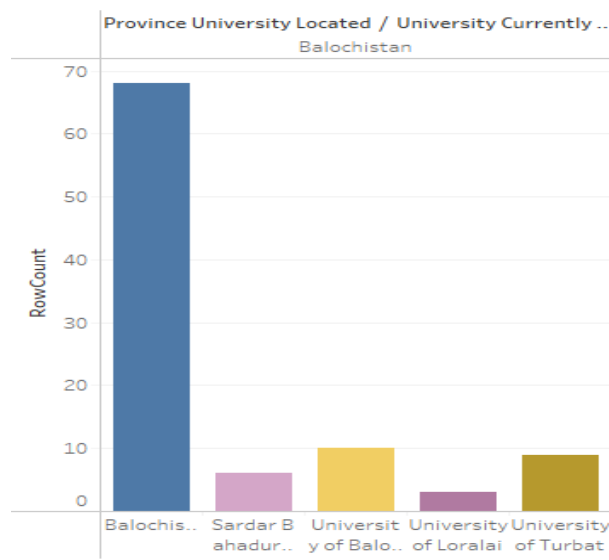
RowCount for each University Currently Teaching broken down by Province University Located. Color shows details about University Currently Teaching. The view is filtered on Province University Located, which keeps Capital.

University Currently Teaching

- Air University
- Bahria University, Islamabad
- Capital University of Science and Technology
- COMSATS, Islamabad Campus
- FAST-NU(Islamabad)
- International Islamic University, Islamabad
- National University of Modern Languages
- National University of Sciences and Technology
- Pakistan Institute of Engineering and Applied Sciences
- Quaid e Azam University Islamabad
- RIPHAH International

In Islamabad, as seen in the chart above, COMSATS university is leading with respect to the count of teachers as compared to other universities overall, and also within the Capital, Islamabad.

balochistan teacher count



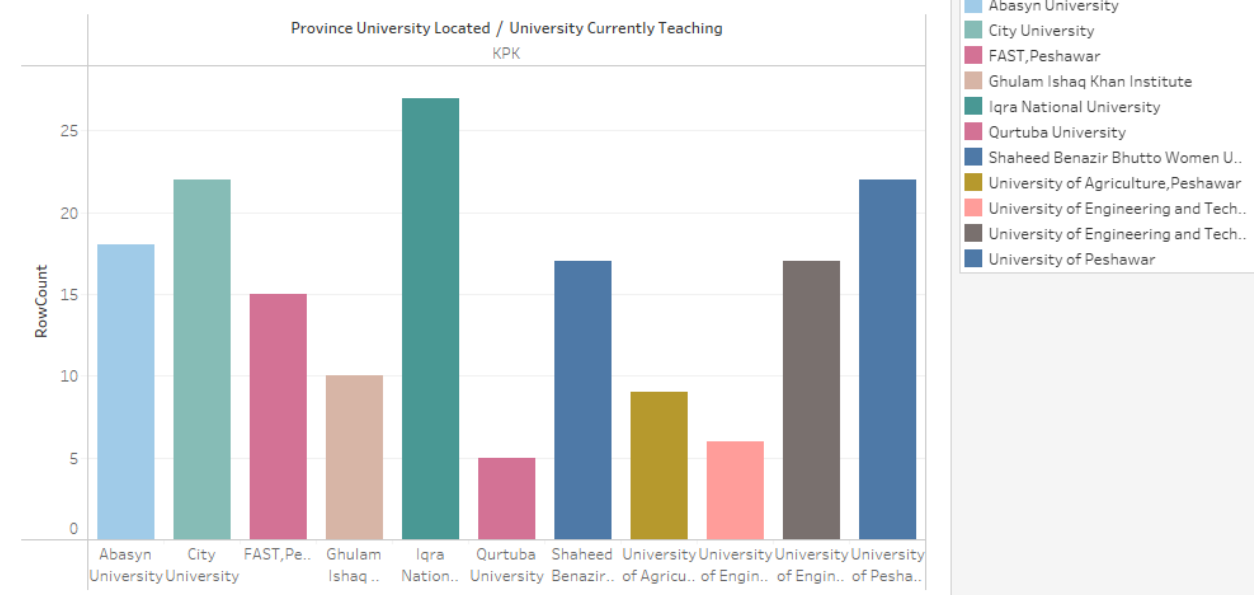
RowCount for each University Currently Teaching broken down by Province University Located. Color shows details about University Currently Teaching. The view is filtered on Province University Located, which keeps Balochistan.

University Currently Teaching

- Balochistan University of Information Technology, Engineering and Management Sciences
- Sardar Bahadur Khan Women's University
- University of Balochistan
- University of Loralai
- University of Turbat

Balochistan has the least number of teachers because of a lack of an education system and thus, lack of universities. However, there seems to be one university that is leading all others by a large margin i.e., the “Balochistan University of IT, Engineering and Management Sciences”. Since this is a CS based university, it has more teachers than all others combined.

KPK teacher count

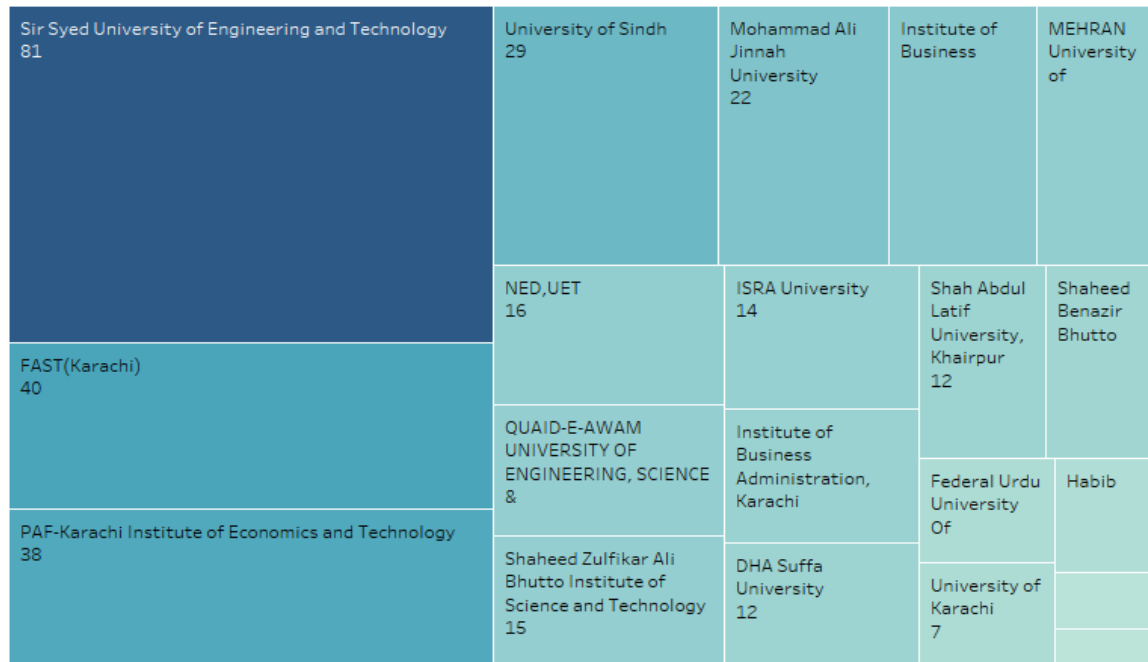


University Currently Teaching

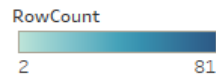
- Abasyn University
- City University
- FAST, Peshawar
- Ghulam Ishaq Khan Institute
- Iqra National University
- Qurtuba University
- Shaheed Benazir Bhutto Women U..
- University of Agriculture, Peshawar
- University of Engineering and Tech..
- University of Engineering and Tech..
- University of Peshawar

The number of teachers in KPK seem to be similar across universities, however, upon noticing the row count, we can observe that the highest count of teachers in KPK is a little above 25, whereas, in all other provinces, the count of teachers is well above 60. This shows the lack of CS studies in KPK.

sindh teacher count



University Currently Teaching and RowCount. Color shows RowCount. Size shows RowCount. The marks are labeled by University Currently Teaching and RowCount. The data is filtered on Province University Located, which keeps Sindh.



In Sindh, Sir Syed University, FAST and PAF are known to have the highest count of teachers, with Sir Syed University having double the count of FAST.

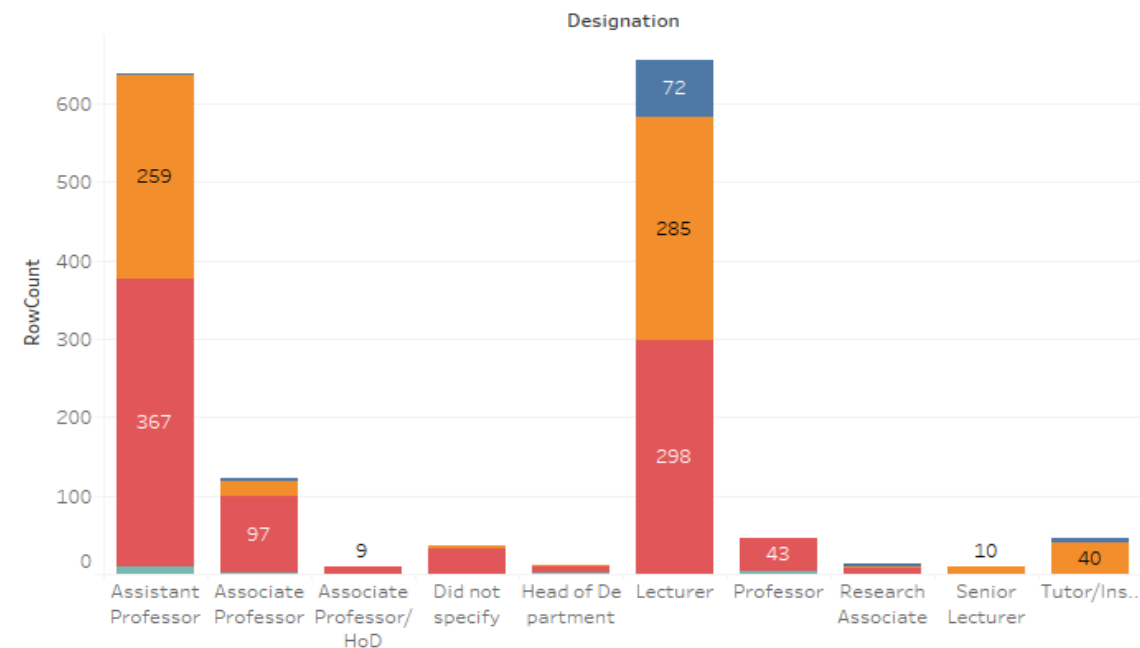
Along with Punjab, Sindh also has a variety of universities with CS departments.

- 2. If Terminal Degree is narrowed down to basic degrees, what are the different types of degrees in each? Comparison of Designation and Terminal Degree to observe patterns between what degree the teacher has done and what they are doing currently.



The dataset contains multiple ways of showing a Bachelor's degree, such as BS which is what 78 teachers have, as their terminal degree. Other Bachelors degrees include BS, BE, BSCS, BIT etc. These are all correct ways of showcasing a degree, so I decided not to remove this column despite having created a new column 'Basic Degree'. The same goes for the terminal degrees related to Masters, PhD and Post Doctorate. Master's degrees have the most variety out of all with MS graduates leading with 469.

designation vs degree



RowCount for each Designation. Color shows details about Basic Degree. The marks are labeled by RowCount. The view is filtered on Designation, which keeps 10 of 57 members.

Basic Degree

- Bachelors
- Masters
- PhD
- Post Doctorate

It can be observed from the graph above the designations that have been chosen by teachers with a certain degree. This graph only contains the top 10 designations, by count of teachers. Teachers with PhD degrees seem to be more prominent from this graph and have more chances with such designations. On the other hand, teachers with Bachelor's degrees are much lesser, with most of them opting to be a Lecturer.

❖ *Note: Final query no. 3 has been explained in Story 3 of my tableau sheet.*

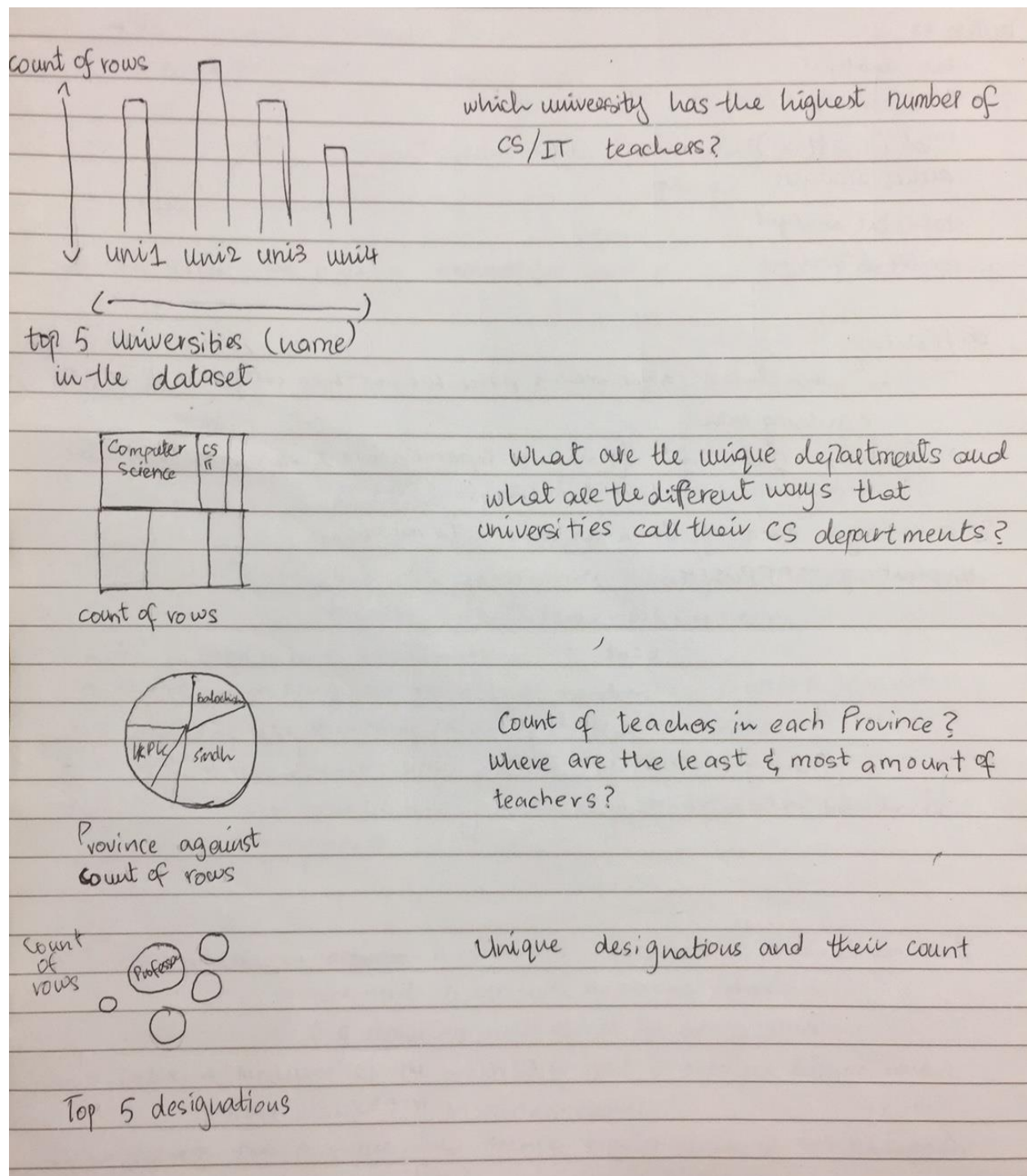
INSIGHTS

- Overall, Punjab has the greatest number of universities, and also the highest count of teachers, exceedingly well over 60. However, COMSATS university in Islamabad has been found to have the highest number of CS teachers (113) in all of Pakistan. Balochistan has only one university for CS, or so we assume according to the dataset, and that is the one with the highest number of teachers, also exceeding 60. KPK has an alarming situation of only 25 CS teachers at the highest in a university.
- Teachers with PhD's seem to be prominent throughout the dataset, which can mean that most universities prefer teachers holding PhD degrees, compared to Masters or Post Doctorate. BS and MS are most dominant amongst other degrees acquired by teachers.
- These statistics show that teachers have majorly graduated from Pakistan with PhD degrees. The count of teachers graduated from multiple countries abroad is much lesser compared to the ones who graduated from Pakistan. This could indicate that universities do not prefer international graduates over local graduates, which is what many assume.
- Upon splitting the column '*Teacher Name*' and producing the '*Gender*' column, apart from the Dr., Engr., and others, I figured out the count of female and male teachers in the current dataset and I found out that there were 104 female teachers, and 120 male teachers. This number is not exact since that would require a deep analysis on the names of the teachers, however, it is an estimate of the ratio of male to female teachers.

SOURCES

- <https://www.kaggle.com/datasets/zusmani/pakistanintellectualcapitalcs> - Kaggle dataset
- <https://bexpress.com.pk/2021/11/the-education-system-in-balochistan/>
- <https://hec.gov.pk/english/universities/Pages/AJK/UniversitiesStatistics.aspx>

IMAGES (SKETCHES)



Dated:

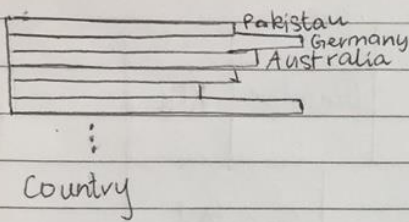
count
of
rows



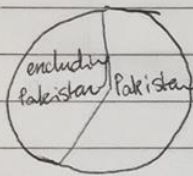
Education status of teachers;
which degree is most found?

Top 5 degrees

count
of
rows

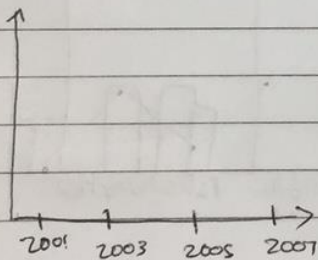


Teachers graduated from
each country & their count



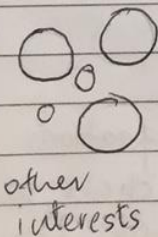
Ratio of teachers graduated
from Pakistan vs abroad

count
of
rows



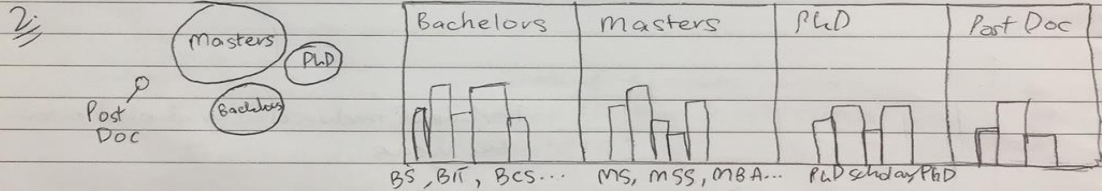
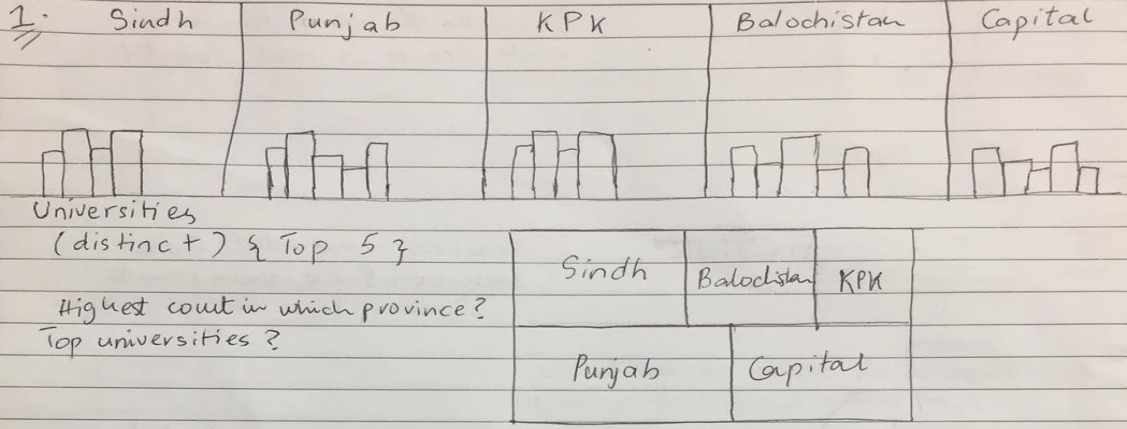
Timeline of teachers who graduated over
the years

count
of
rows



What are the different types of
extra information and what is most
dominant?

Dated:



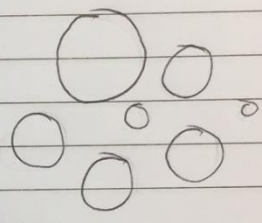
Designation
1		

Break down how many designations there are and use bar chart to show patterns between Basic Degree and Designation

Dated:

3. Use map if possible to show the countries that teachers graduated from.

country	country	country
country	country	..
...
...



Countries other than Pakistan - Analyse