

# DATA 512: Human Centered Data Science

## Project Report

### Analysis of COVID-19 Vaccination's Effect on Employment in Salt Lake County, Utah

## Introduction

Apart from the devastating loss of life that the pandemic caused to society, it also upended the lives and livelihoods for many of us. Due to the contagious nature of the disease, society needed to limit social interactions which had many ripple effects on people ranging from leaving them unemployed to affecting their mental health.

I chose to explore the impact that the pandemic had on employment, and see which sectors it affected and how, what did their recovery look like and what role did vaccinations play in that recovery. Specifically, I chose to look at multiple sectors because my sample of “thick data” of COVID-19 experiences is limited because I mostly interact in circles where work shifted to a remote culture. The reason I chose to incorporate vaccination data is because much of the slowdown in the spread of COVID-19 that I observed in Part 1 of the project seemed to come after the introduction of vaccinations.

It is important to look at lived experiences that are different from ours and recognize that the pandemic was a vastly different experience for each one of us, and it is important to be aware and maintain curiosity of what this experience looked like for other people, as best as possible using whatever means we have at our disposal.

## Background/Related Work

My analysis and research questions are mostly motivated by the fact that in India (where I was living when the pandemic hit), we had extensive coverage of how daily wage workers that migrated from their villages to work in big cities were most adversely affected by the pandemic. They were left abandoned by the government for the most part and could not provide for themselves having lost their only source of income with any effort from the government arriving too late or still inadequate to support everyone who was affected. There were reports of how many of them elected to go back to their homes which were thousands of kilometers away on foot; unfortunately, many who made this decision lost their lives, even the ones who did were not sure what they could do next.

An excerpt from an article [1] paints a clearer picture of how dire the situation was:

*“All transport—roadways, airways and railways—were suspended, including hospitality industries, educational institutions, and industrial units. As the factories and workplaces closed down, millions of migrant workers had to deal with loss of income, food shortages and an uncertain future. The scale of this issue varies from state to state or city to city, but has caused widespread disruption. With no money, no job, unsure when the lockdown will finally end, the migrant workers had no other option than to return back to their villages. Their massive migration*

*from working states has formed a humanitarian and health security challenge and an exceptional logistical nightmare.*

*This instigated the next problem for them. How do they reach home? With road and rail transport links still suspended, walking back was the only option and they initially took to the road. Images of marching migrant workers most of them left with nothing but keen to reunite with their families back home. People have undertaken hazardous journeys, sometimes walking up to 1000 km with no money to spend and often without food for days together. Many were arrested by law enforcement officials for violating the lockdown, some died due to exhaustion or accidents on the roads.”*

BBC also has a video piece [2] on the migrant crisis and illustrates the difficulties migrants faced and how they were paid less even if they were fortunate enough to find work amidst all the uncertainty and negligence they endured.

Having seen how drastic the effect of the pandemic was on some members of our society, I wanted to investigate if there were similar patterns in the US, or if there was something contrary to the assumptions that I held. I chose to relate these findings with the vaccination timeline, as the US was quick to roll out vaccinations even as it floundered in the beginning with other pandemic containment measures. Vaccinations were the promised hope to a life that is largely normal, and a huge part of that is resuming work and being able to make a living. Armed with these thoughts, my research questions were two-folded:

1. How were different employment sectors affected by the pandemic?
2. How did the adoption of vaccinations influence employment?

Each of these questions were investigated to limit the scope of inquiry to Salt Lake County, Utah, the country that was assigned to me for the first part of the analysis.

## Methodology

The primary motivation of this analysis extension was to tell a story about how the situation evolved for different job sectors in Salt Lake County and for that I concluded that this analysis should primarily be supported by exploratory visual analysis of the data. A lot of iterating over different visualizations would be needed so that the story about how the pandemic affected employment could be told in a manner that does not bias the viewer by reflecting my assumptions or preconceived notions about what we can expect.

Additionally, there are many benefits of data visualizations. We can encode much more information through a single visualization, focus on the bigger picture, i.e., the story at hand rather than find ourselves fixated on the assumptions that statistical models and testing require as part of due diligence before we apply any methods. Additionally, it is vital to note that many results that prove statistically significant are perhaps not practically significant. A low p-value merely indicates to us that the distributions are different and offer no indication of how meaningful the observed difference might be.

We are most interested in understanding the time span that the employment situation took to improve on a month-over-month basis, what that improvement was on a year-over-year basis to account for seasonal effects, and how these observations dovetail the coverage of vaccination against COVID-19. We outline our process in the steps below:

1. Project Part 1
  - a. Get COVID-19 daily case data and mask mandate enforcement data
  - b. Remove noise from the cases data by taking a moving average
  - c. Calculate the day-over-day percentage change
  - d. Plot the rate of change of COVID-19 cases and highlight time periods of different states of mask mandates to observe if there is a correlation
2. Project Part 2
  - a. Collect data from BLS for employment and CDC for vaccinations for the county, largest city or state level data based on availability, in that order of preference
  - b. Calculate month-over-month and year-over-year percentage changes for each sector's employment counts
  - c. Identify the following points in the vaccination data – date when vaccinations began, date when 25%, 50% and 80% of the population that is 18 years old and above and is vaccinated
  - d. Plot the raw employment counts, percentage changes on a monthly and yearly basis as time series and mark vertical lines on the plots for the dates identified in the previous step
  - e. Identify if any dips or spikes occur before or during the vaccination phase. We can attribute these observations are correlations but not causation

## Findings

The spread of COVID-19 seemed to slow down after the introduction of vaccinations and not so much due to mask mandates

This conclusion comes from Part 1 of the project where we plotted the growth rate of COVID-19 cases and tried to correlate it with mask mandate policies enforced at different times.

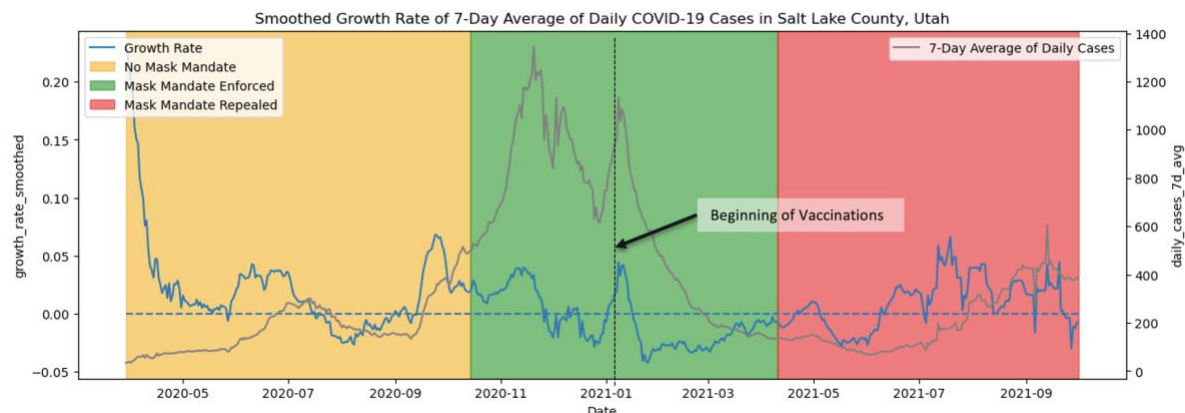


Figure 1: Part 1 of the project relating growth rate of COVID-19 cases to masking mandates

We can see that the period where growth rate is mostly below 0 is after the point when vaccines were available in the county. Although there is a dip in cases before this too, it could be due to a delayed effect of effective masking and mask quality, or just because more people chose to stay at home.

Many sectors made good recoveries even before vaccination was available

We saw that overall, and for many sectors, the employment counts had reached back to or surpassed 2019 levels in a matter of months. Here are some sectors that bounced back quickly:

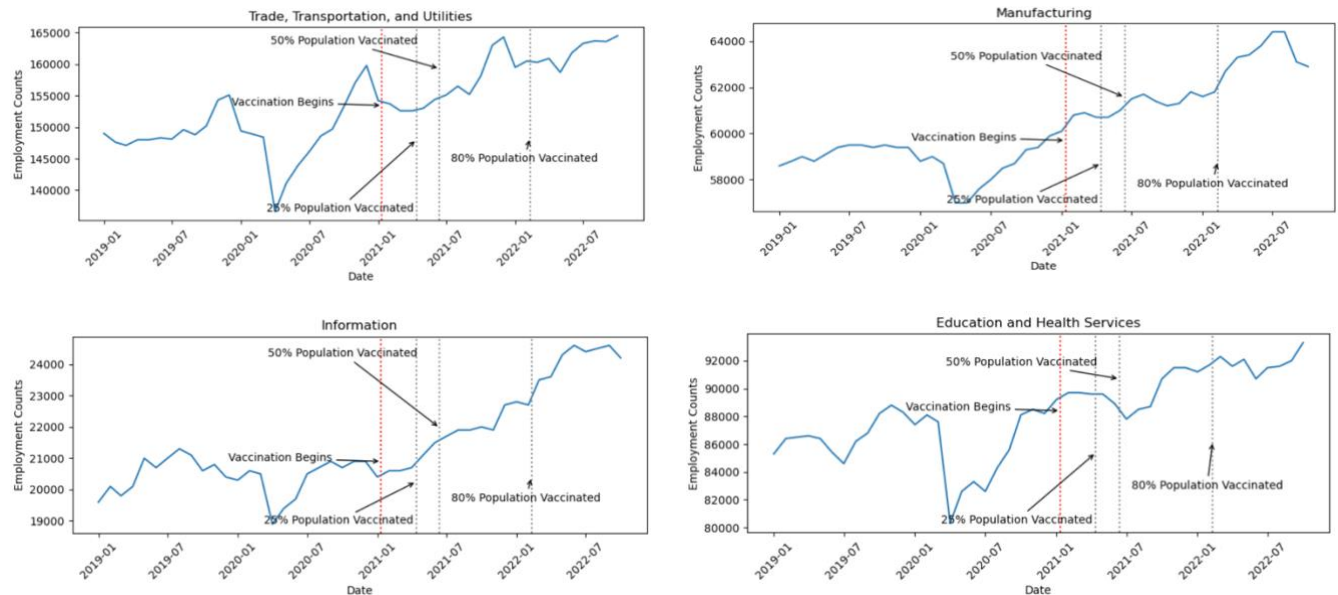


Figure 2: Sectors that made recoveries before the introduction of vaccinations

And then there were sectors that did not see a substantial decrease in employment immediately as the pandemic struck. The financial and mining sectors have grown with much smaller disruptions for this county.

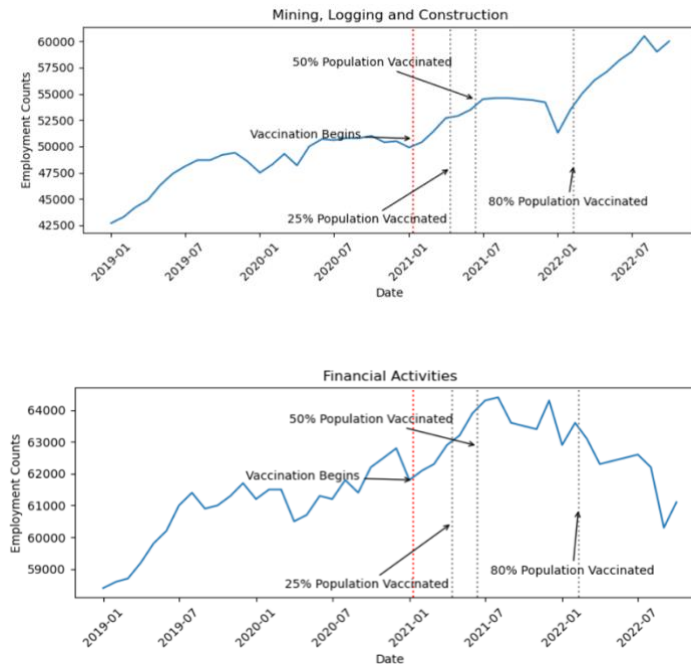


Figure 3: Time Series for Employment Counts for Financial Activities and Mining, Logging and Construction Sectors

Although, there were exceptions. The Leisure and Hospitality sector was worst hit and took almost 2 years to recover close to 2019 numbers. The Government sector has a similar count for 2020 and 2021, both are below 2019 numbers and do not look to be reaching those numbers in the short run, based on the plot we have observed so far.

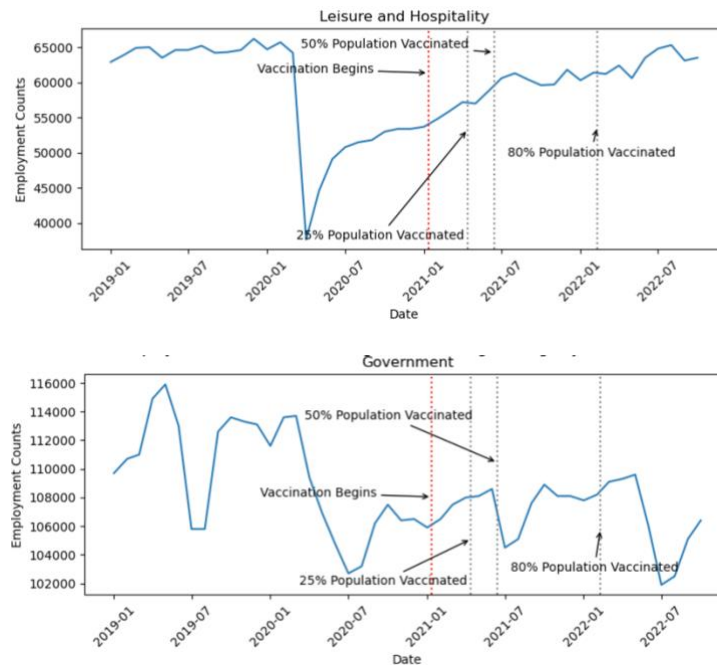


Figure 4: Time Series for Employment Counts for Government and Leisure and Hospitality Sectors

There was a spike in employment for many sectors even before 25% of the population was fully vaccinated

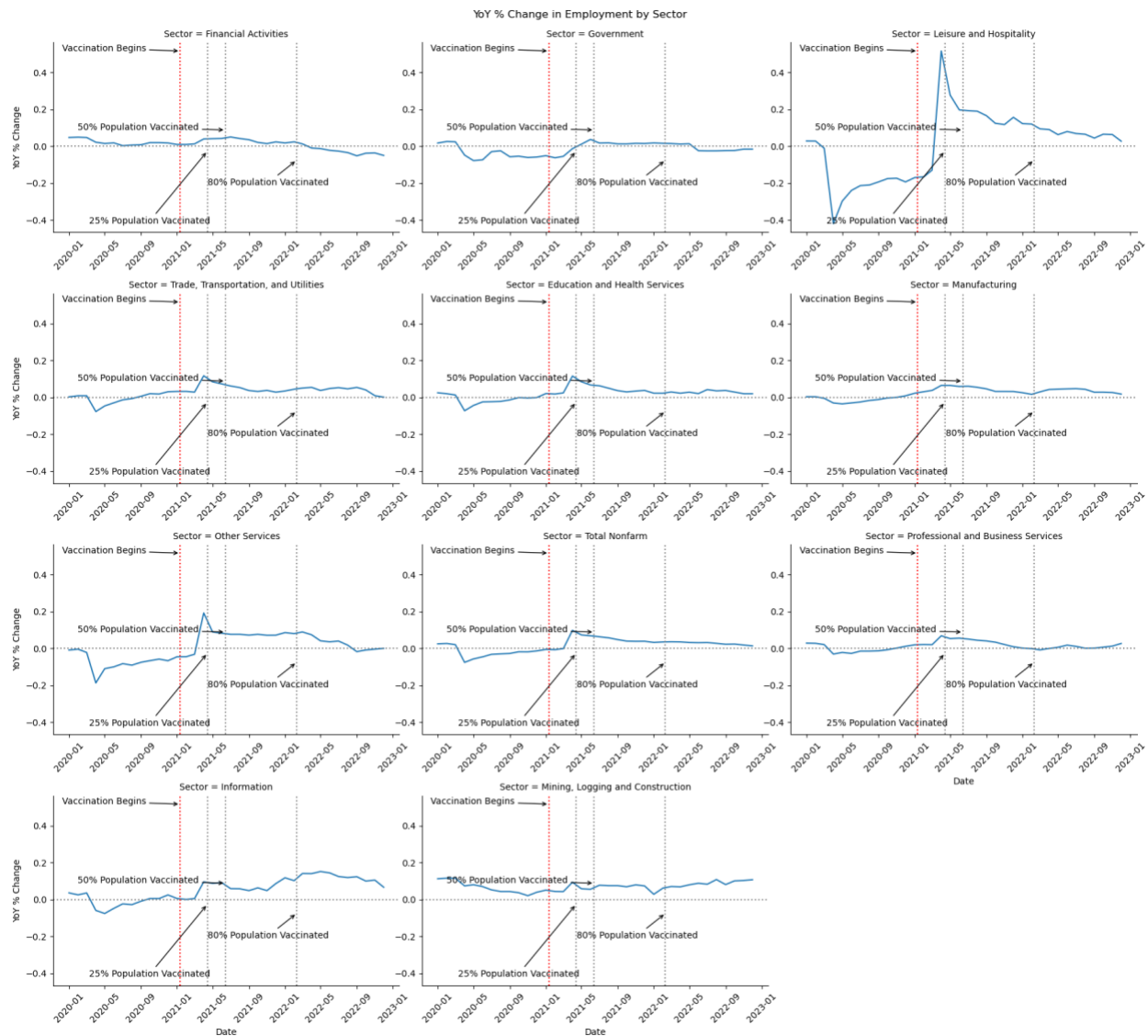


Figure 5: Year-over-year percentage change in employment by sector

In addition to the observation that we see in the sub-heading, we can also confirm many of the notions we had about employment counts from this graph. We see that there is a spike in employment for most sectors right before the 25% vaccination mark. This is interesting because much of the disruption that daily life faced was expected to be maintained in some form until a sizeable population was fully vaccinated. This could either be a premature faith in vaccinations, or a sudden change in workplace and social distancing regulations that encouraged employers to increase hiring.

## Discussion/Implications

First thing that we can conclude from this analysis is that while there was a negative effect of the pandemic on employment for this region, the situation was not as dire as that observed in India, which inspired this research question to be explored. We tried to cover a considerable breadth of industries so that we can have a representative sample of different sections of the population, even as the data was not pinpoint enough to our liking, which we discuss in the limitations section.

Through this analysis, we were able to identify that not all sectors were affected to the same extent by the pandemic and not all sectors had a similar recovery story either. It would be interesting to understand if the story was similar for other counties or whether there were substantially different patterns that emerged because of difference in policy decisions and economic needs of a county.

It is noteworthy that even before less than 25% of the population was fully vaccinated, there was a huge uptick in the growth of employment on a year-over-year comparison. In the future, a question that can be explored is whether this uptick was premature and if it caused more harm than good, in terms of containing the spread of the disease. This question is pertinent because, experts estimated that to build herd immunity we would require upwards of 70% of the population to be immune [3].

The simplistic nature of our analysis method helped us understand at a high level not only the story of how the pandemic affected different employment sectors, but also ask follow-up questions such as why we saw some things and whether these observations were in line with guidelines recommended by many epidemiological experts to fight the pandemic

## Limitations

Some of the limitations of this study stem from the data that was collected. While it was helpful to see how different sectors were affected, the groupings that are created by BLS are still not very granular. A lot of occupations that are grouped into a single sector probably adapted differently to the pandemic. Some examples:

1. Leisure and hospitality: Different sub-sections of this sector, such as hotels, restaurants and tourism could have been affected differently and perhaps had different timelines depending how restrictions imposed by the pandemic affected each of these sub-sections
2. Education and health services: These two sectors were dealt with in a completely different manner. Healthcare was an essential service and was particularly strained during the pandemic, whereas education at one point almost came to a complete standstill before it was deemed possible to resume learning and eventually school attendance
3. Other services: Many occupations are clubbed under this category. Here is the definition from BLS –



*“The Other Services (except Public Administration) sector comprises establishments engaged in providing services not specifically provided for elsewhere in the classification system. Establishments in this sector are primarily engaged in activities, such as equipment and machinery repairing, promoting or administering religious activities, grantmaking, advocacy, and providing dry-cleaning and laundry services, personal care services, death care services, pet care services, photofinishing services, temporary parking services, and dating services.”*

Having more detail about sub-sections of this sector would be most interesting as we would be able to see if there were any major shifts in occupations that were not as widespread as the other categories.

From these examples it is clear that the data is not perfectly suited for the analysis we wished to conduct as it was not collected for this specific purpose by the BLS. Additionally, we could not find data that would have been insightful in this analysis. For instance, we did not find a demographic breakdown of these occupations or income distribution for these occupations. We could have perhaps observed discrepancies in how different demographics were affected by the pandemic or investigated if the standard of income was commensurate to years before the pandemic, even as employment counts bounced back quickly.

Other considerations that we must make are that there are also other macro and microeconomic factors at play that influence employment such as business cycles (which are not perfectly seasonal) or local legislation, and we must be careful to attribute everything we may observe to the pandemic alone. Many of these effects often interact with each other and have complexity to them. This is precisely why we have not pursued causal modeling or regression analysis – because such analysis requires accounting for every possible confounder that we are able to identify. This analysis can merely be limited to as correlation.

Lastly from a data standpoint, any results that we observe can only be concluded as applicable for Salt Lake County. Economic activity and sociological setup of every region vary heavily and the findings for this region cannot be assumed for other regions.

## Conclusion

From this study of employment and vaccination data we were able to make some interesting observations in employment patterns with respect to the pandemic timeline for many sectors and identify which sectors were hit hard and which ones kept surging despite how the pandemic unfolded. We found that while there was a huge impact on employment, it did not seem to compare to be as dire as that in India and the recovery was quick, for the most part. We also found that there were spikes in employment growth a bit early into the vaccination phase of the pandemic – many people got back to work even though we had not immunized a sizeable population against the disease. Determining whether this was reckless or not from a public health standpoint can be taken up as an analysis in and of itself by approximating what fraction of the population was susceptible to infection at different points of the pandemic.



Understanding the timeline of employment recovery helps us build intuition of the ideas that guided policy in Salt Lake County, Utah, and the tradeoffs the administration was willing to make. To build a complete understanding of this, we must also understand the timeline from a “thick data” point of view and not just make assumptions off data that we were able to analyze and observe. Qualitative measures of how the workplace evolved are not documented in this analysis – knowing a little more about them would be valuable to understand the administration’s point of view. Though we mustn’t forget, a quantitative analysis of the data is what led us to identify this issue to be of pertinence in the first place.

## References

- [1] Iyengar KP, Jain VK COVID-19 and the plight of migrants in India, *Postgraduate Medical Journal*, 2021;97:471-472. <http://dx.doi.org/10.1136/postgradmedj-2020-138454>
- [2] India migrant workers paid heaviest price for Covid crisis, *BBC*, December 2020. <https://www.bbc.com/news/av/world-asia-india-55434594>
- [3] “How Much Herd Immunity Is Enough?”, *New York Times*, December 2020. <https://www.nytimes.com/2020/12/24/health/herd-immunity-covid-coronavirus.html>

## Data Sources

Salt Lake City, Utah, Metropolitan Area Data Tables, U.S. Bureau of Labor Statistics – [https://www.bls.gov/regions/mountain-plains/data/employmentandunemploymentandwages\\_saltlakecity\\_table.htm#ro7qcew-saltlakecity5.f.1](https://www.bls.gov/regions/mountain-plains/data/employmentandunemploymentandwages_saltlakecity_table.htm#ro7qcew-saltlakecity5.f.1)

This dataset has a variety of tables in time series format pertaining to employment counts in non-farm occupations, overall and broken down by industry super sectors

COVID-19 Vaccination Data, Center for Disease Control - <https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh/data>

Consists of a time series of COVID-19 vaccination rate by county, updated at a weekly cadence. We use this data to mark different points of vaccine adherence and see how they are related to other patterns we may observe in the data