

IT Center Group Project

Yitong Tang, Immanuel Kraft & Alexander Atzberger

30 January 2020

Introduction of the data

We examined the employee behaviour of a Chicago based IT Call Center office. 28 employees at the firm were examined, of which 23 participated in the study. Nineteen-hundred hours of data were collected, with a median of 80 hours per employee. The resulting data document the performance of computer system configuration tasks assigned to employees on a first-come, first-served basis. These configurations were rated to one of three levels of difficulty (basic, complex, or advanced) based on the configuration characteristics. At the conclusion of the task, the employee submitted the completed configuration as well as the price back to the salesman, after which the employee moved to the back of the queue for task assignment.

To conduct analysis on the data set, we decided to zoom in on the week before and right after the Easter holiday break in a time frame from 09:00-17:00. All subjects across all three departments (Pricing, Configuration & Management) were utilized, examining their task behaviour, with closely examining the closing time, complexity and movement across departments between the two weeks. Questions about the generalizability arise when examining the data: only two weeks among employees were compared, there are potentially more confounding variables that influence the outcomes of the presented visualizations.

In the project following R packages were used: **tidyverse** (Wickham, 2019), **lubridate** (Spinu, Grolemond, & Wickham, 2018), **dplyr** (Wickham et al., 2019), **citr** (Aust, 2019), **bibtex** (Francois, 2020), **plotly** (Sievert et al., 2019), **igraph** (file., 2019), **networkD3** (Allaire, Gandrud, Russell, & Yetman, 2017) and **prettydoc** (Qiu, 2019).



The layout of the workspace is shown in the figure above. Participating employees are indicated at their cubicles by their badge IDs; different colors behind the IDs represent different departmental branches at the firm. Non-participating employees have letter "N" at their cubicles. Employees fetched their badges from the room containing base station 1 (located at the lower left corner) at approximately 9am each weekday morning, and returned the badges to this room at around 6pm in the evening. The RSSI regions were manually assigned to identify different regions in the workspace, and do not correspond to any particular sensors deployed in this experiment.

Tidy Data

In order to conduct a meaningful analysis, we tidied the data. The core of the analysis of the data is to compare the week before the Easter Break 2007 with the week right after, selecting observations from **02.-06 and April 2007 - 23. - 27. April 2007**. In a subsequent step, we separated the `date.time` variable into separate `date` and `time` variables and created two separate new data frames for the `IR`, `Transaction` and `Location` data frames, resulting in a total of *six* explored data frames. In the final step, we removed all the missing values and duplicate values from the active datasets.

Data Frames & Primary Keys

Location

The dataset location describes the coordinates of the workers and their position in the office at a certain point in time and how they move across the building. The dataset tracks the movement every minute and shows the location. **There is no primary key in this data set because an employee might stay at a location for longer than one minute.**

Transaction

The dataset transaction shows information on the individual employee, including what job they have (configuration/pricing), the task complexity, the amount of follow ups and how long they took to close a task. There is one primary key: the combination of the variables `transact + employee + phase` forms a primary key.

IR

The IR dataset indicates when an employee is close to another employee and/or an anchor node. As it shows the movement track of one ID in every minute. in every exact time, one sender is contacting with an only person. Thus we conclude that we identified the primary key: `sender.id + date + time`.

Primary Keys for Transaction and IR

```
## [1] 2
```

```
## [1] 2
```

```
## [1] 0
```

```
## [1] 0
```

```
## [1] 0
```

```
## [1] 0
```

Data Joins & Foreign Keys

We identified the foreign keys in the datasets:

1. `employee, date2, time2` in `Transaction_clean_first`
2. `employee, date2, time2` in `Transaction_clean_second`
3. `id, date2, time2` in `location_first_week`
4. `id, date2, time2` in `location_second_week`
5. `sender.id, date, time (or local.id, date, time)` in `IR_first_week_j`
6. `sender.id, date, time (or local.id, date, time)` in `IR_second_week_j`

Data Joins & Relationships between tables

Enclosed we display the verification process of the relationships in the data. In order to examine the relationships, we checked the joint key first, to see whether the joint key is duplicated. In our data frames, all of our joint keys we utilized appeared more than once, creating a “many-to” relationship. Subsequently, we utilized the `anti_join` function to

check the number of mismatches and duplicates. The duplication of keys indicates the “to-many” relation of the data. Our relations between tables are displayed below.

1. The relation between Transaction_clean_first and location_first_week is many-to-many
2. the relation between Transaction_clean_second and location_second_week is many-to-many
3. the relation between IR_first_week_j and location_first_week is many-to-many
4. the relation between IR_second_week_j and location_second_week is many-to-many

Floorplan



Interaction between two specific employees in one week

by Yitong Tang

In this plot, I want to explore the interaction between two specific employees in one week. The width of the links between two points show the frequency of the interaction, the colour of the points shows the roles of them, and the name on the points show each employee's ID.

Conclusion

if we check the data set by role: the employees conducting the same row have more interaction with each other: we can find the green points are crowded together, and the light blue points(configuration) are crowded together. the employees in the role of

From the perspective of every specific employee, we can find the most frequent interactions happen between employee 292 and 263. Different from many employees who only interact with the people in the same roles. 292 and 263 interact with various roles. so we can suppose 292 and 263 are gathering information from their group members and transfer the idea to the other role. On the other hand, several employees seldom contact with the colleague: 109 106 253 261. all of them have only one link with the group. However, I think the link is very important because the point they link with is an information-gathering point. For instance, 261 connect with 292, which is an information centre; 253 connect with 290, which also links with many other employees.

[illegible]

by Immanuel Kraft

file:///Users/I507303/Library/Mobile Documents/com~apple~CloudDocs/Documents/UvA/Masters/Data Wrangling R/R course/MasterDirectory_Alex 3/ITCen... 5/11

Companies often have some deadlines or projects to complete before the holidays. Therefore the movement before Easter as well as after Easter was considered and if there are significant differences between the two departments. The graph should also show if there are differences in the number of follow ups, which means how often people exchange before they close an task.

To show this, an animated scatter diagram was chosen, which shows the different movement patterns from Monday to Friday. The x- and y-axis show the coordinates of the office, the color differs between the departments "Configuration" and "Pricing", the size of the points shows the frequency of the follow ups.

First Week

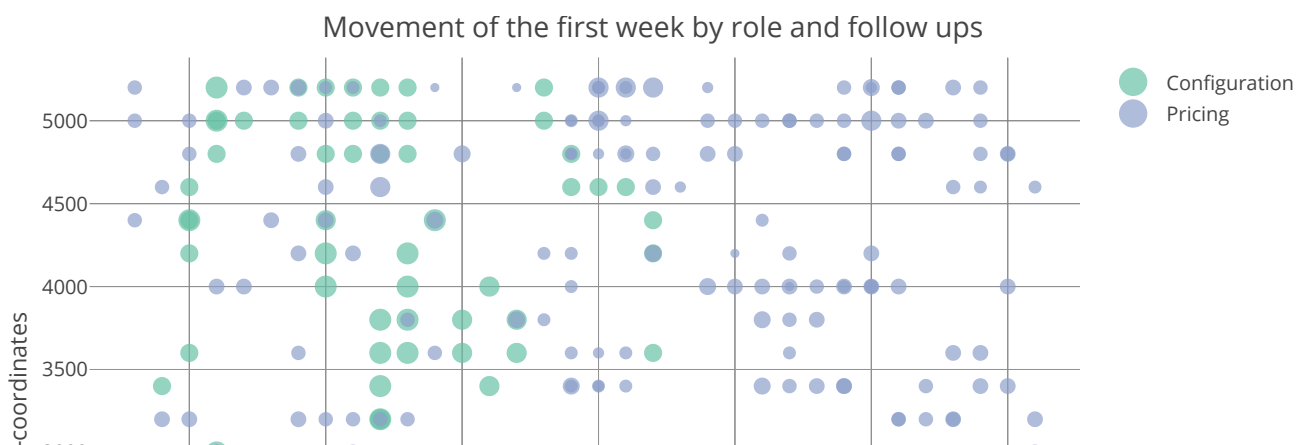
The main results are that during the first week people working in Configuration hardly move to other departments, while Pricing moves throughout the office. If there is an exchange between departments, it always takes place in the Configuration Department or in the meeting rooms below. If you look at the follow-ups, they are evenly distributed between the departments. It is also interesting to note that on the last day of the week the Pricing Department was most likely already off duty, as there is no record of movement.

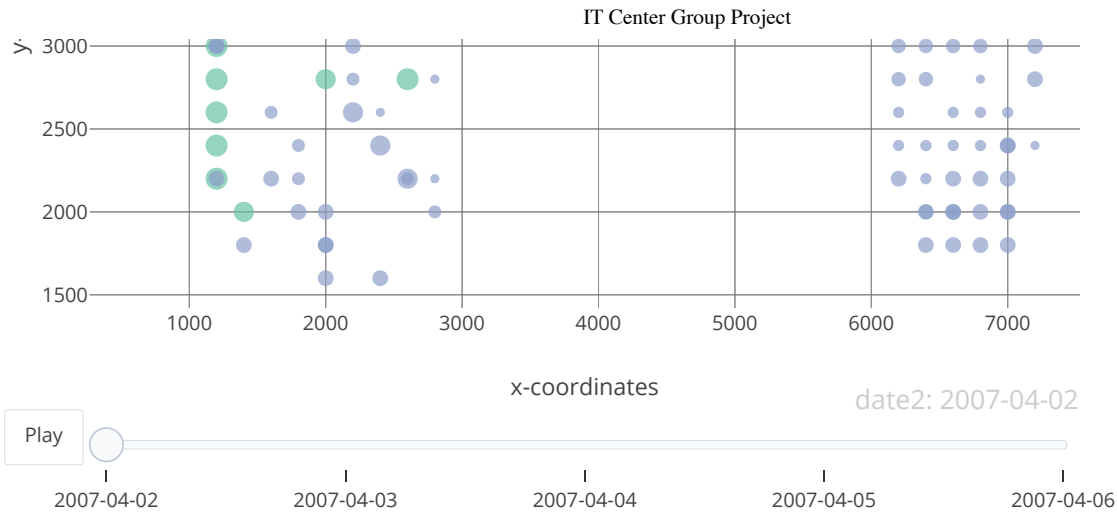
Second Week

Looking at the second graph, it is noticeable that there is less movement overall than in the first week. Also the Pricing Department shows no movement at all, which might be related to the fact that the department was closed for holidays. The Configuration Department also moves only in its own "zone". Apart from the last day of the week, they are never in a meeting room at any time.

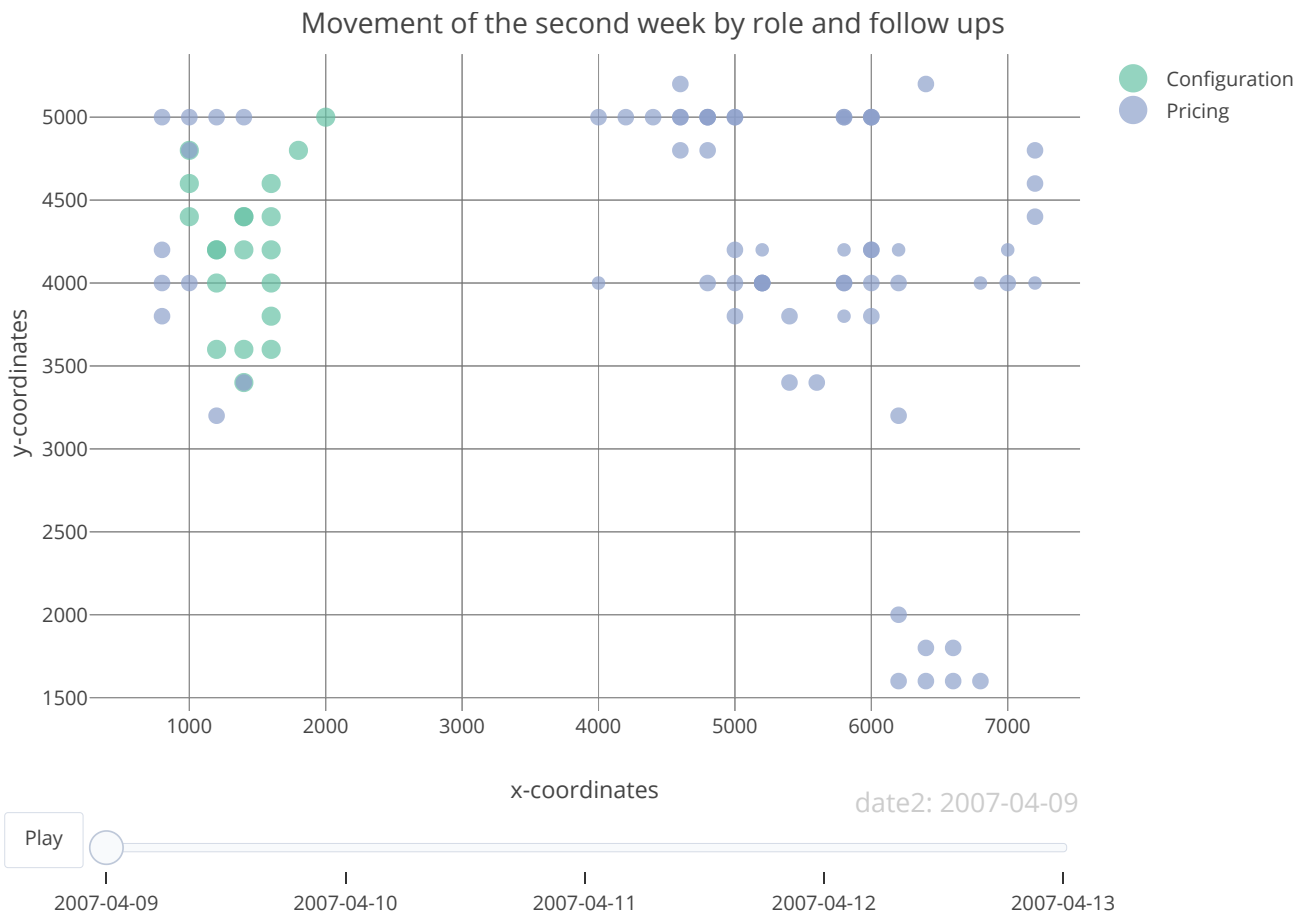
The observation suggests that there was actually more exchange between the employees before Easter and that meetings were scheduled more often in the first week. Since there was less movement during the second week one might conclude, that employees have more stress before public holidays and they try to finish their projects. Another interesting finding is, that the pricing team moves all over the office, while the configuration team stays in their own department. The Management could think whether it is still worthwhile to have permanent working tables for the pricing employees or if a more open approach would make more sense.

Week 1





Week 2



Tracking the movement of the most productive employees across two weeks

by Alexander M. Atzberger

A key psychological hypothesis behind organizational theory is transactive memory: an organization coping with complex tasks often needs a knowledge repertoire far beyond the memory capacity and reliability of any individual in this organization. Individuals collaborate to store this total repertoire by identifying the expertise of one another and distributing the repertoire among themselves. The employees indicated that the configuration and pricing tasks were information-intensive, and therefore required them

to talk to one another to fully understand the various specifications. Thus, the researchers that provided the data expected a positive correlation between the rate of problem-solving by an employee and the number of places visited by that employee, displaying that employees talk to their colleagues.

In my visualization, I compare the two most productive employees of the two departments *pricing* and *configuration* and compare the relationship between closed calls and which department has moved the most across the office.

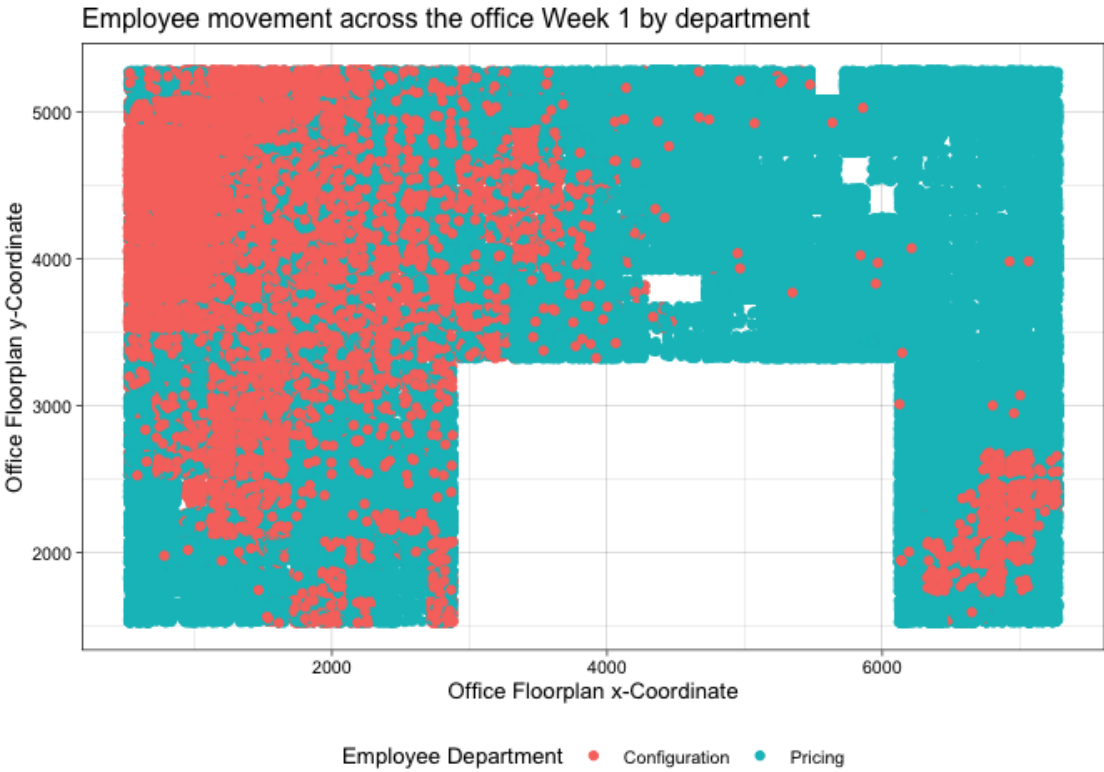
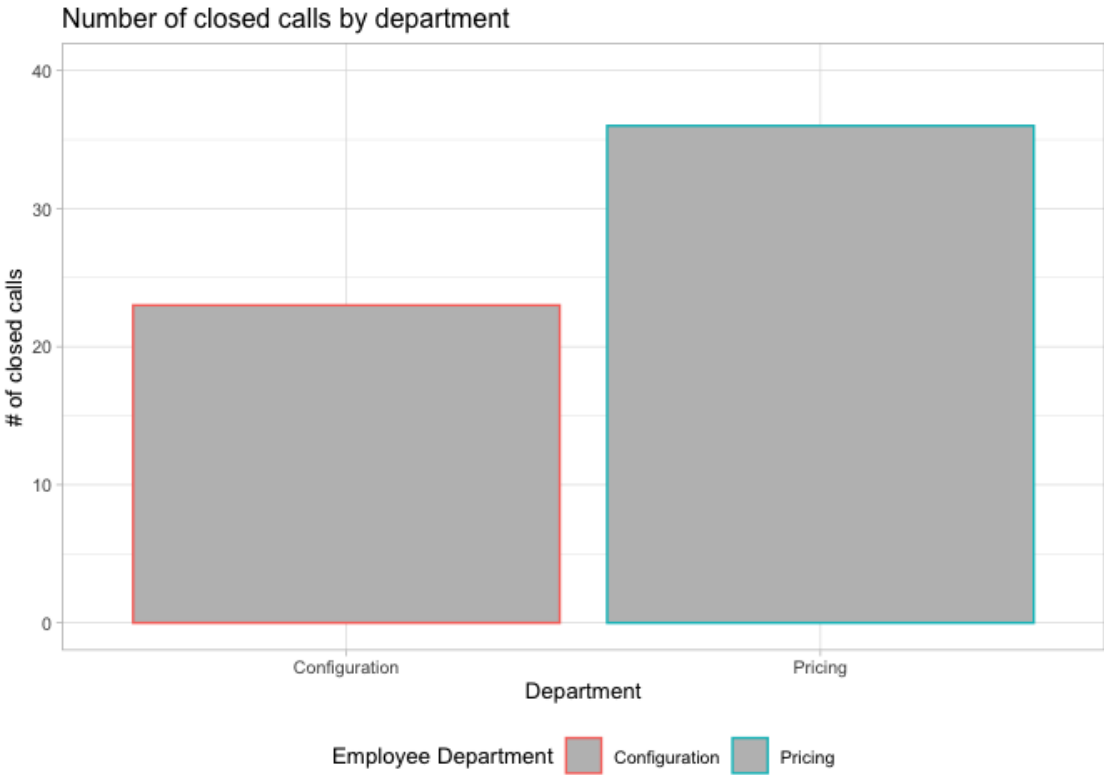
The bar charts implies how many problems the two departments solved in the two compared weeks, with the department on the x-axis and frequency of solved calls on the y axis.

The jitter plots below the bar charts display the pathways that each employee (grouped by department) took in each of the compared weeks. The x and y axis represent the coordinates of the office layout. The two colors represent the two different departments.

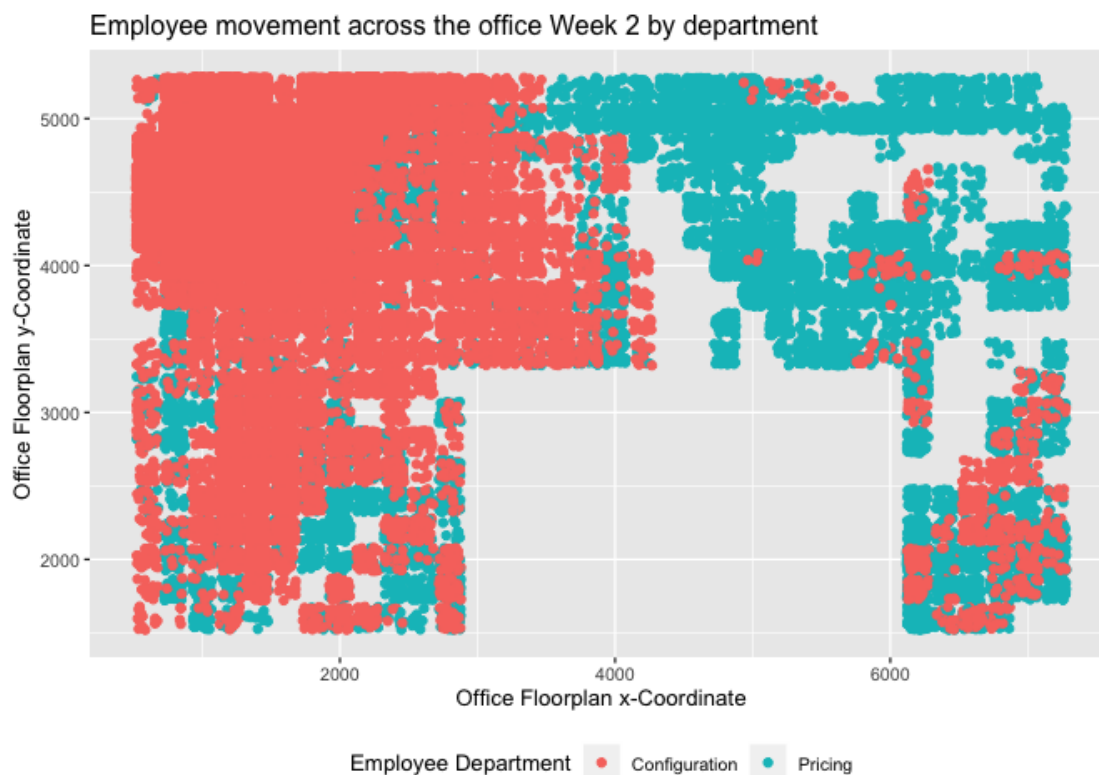
Conclusion

Overall, the employees from the IT Call Center had a more productive week in the week before the Easter break. In Week 1, the *configuration department* closed 23 calls, the pricing department closed 36 calls. In Week 2, the *configuration department* closed 15 calls, the pricing department closed 21 calls. The employees of the pricing department close more calls than the configuration department in the first week and move more across the office space. Interestingly, in the second week, the employees of the pricing department again closes more calls during the week, but move significantly less across the office. This finding challenges the researchers initial assumption of the positive correlation between task closing and movement across the office. Regardless, the results of the visualization need to be interpreted with caution, as observations from a two week time frame with a holiday in between pose difficulties for the external validity of the findings: there are a number of potential correlating factors that potentially explain the results.

Week 1



Week 2



References

Allaire, J., Gandrud, C., Russell, K., & Yetman, C. (2017). *NetworkD3: D3 javascript network graphs from r*. Retrieved from <https://CRAN.R-project.org/package=networkD3>

Aust, F. (2019). *Citr: RStudio add-in to insert markdown citations*. Retrieved from <https://CRAN.R-project.org/package=citr>

file., S. A. (2019). *Igraph: Network analysis and visualization*. Retrieved from <https://CRAN.R-project.org/package=igraph>

Francois, R. (2020). *Bibtex: Bibtex parser*. Retrieved from <https://CRAN.R-project.org/package=bibtex>

Qiu, Y. (2019). *Prettydoc: Creating pretty documents from r markdown*. Retrieved from <https://CRAN.R-project.org/package=prettydoc>

Sievert, C., Parmer, C., Hocking, T., Chamberlain, S., Ram, K., Corvellec, M., & Despouy, P. (2019). *Plotly: Create interactive web graphics via 'plotly.js'*. Retrieved from <https://CRAN.R-project.org/package=plotly>

Spinu, V., Grolemund, G., & Wickham, H. (2018). *Lubridate: Make dealing with dates a little easier*. Retrieved from <https://CRAN.R-project.org/package=lubridate>

Wickham, H. (2019). *Tidyverse: Easily install and load the 'tidyverse'*. Retrieved from <https://CRAN.R-project.org/package=tidyverse>

Wickham, H., François, R., Henry, L., & Müller, K. (2019). *Dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>