

# Rejuvenation of an Old Plot

---

Group 24 - Date: 2022/10/26

## Team Members

Name	Email	Student ID
Chengkai Lin	linchk20@lzu.edu.cn	320200945841
Wenbo Liu	wbliu20@lzu.edu.cn	320200931121
Yifei Guo	guoyf20@lzu.edu.cn	320200930901
Xuda Han	hanxd20@lzu.edu.cn	320200945781

## Notes

- The file was only written and tested using *Visual Studio Code* <sup>Version 1.72.2</sup>, and previewing with other softwares are **not guaranteed** to render correctly. Sorry for the inconvenience if this problem occurs!
- Only **two** images (the *replica* and the *improved plot*) are embedded using *svg* format, since the other images was **originally bitmaps** and converting is unnecessary.

## What We Did

In this project we will choose a problematic plot from the Internet, then try to replicate it, and apply some principles to make it look better and less confusing.

## Information Visualization

Before we start, I would like to make a brief introduction about information visualization. This is the definition we took from [this website](#).

Information visualization is the process of representing data in a visual and meaningful way so that a user can better understand it.

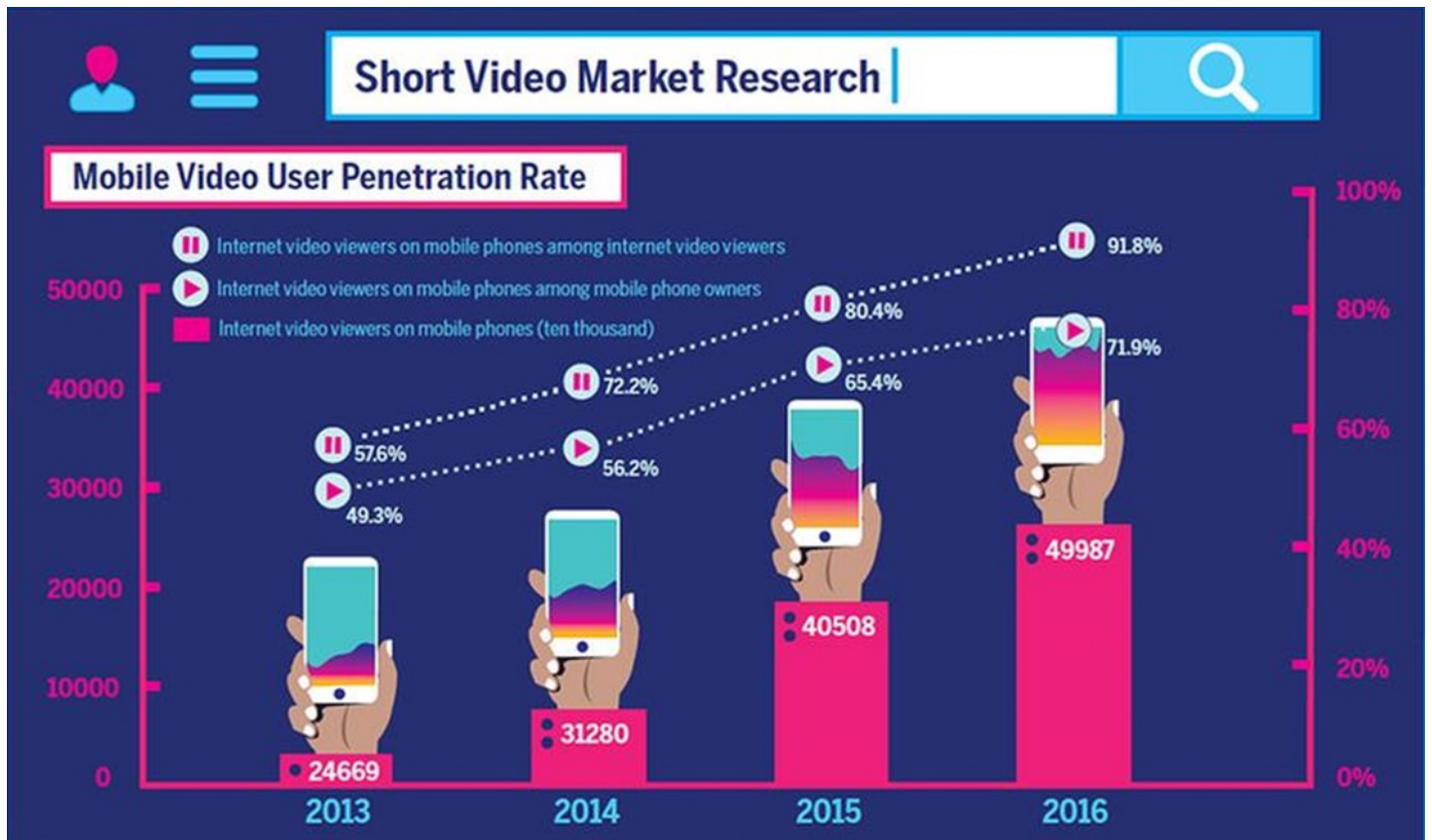
It means that the importance of information visualizaion is that it exploits the hidden values of data and express them in an intuitive and efficient way.

However, making a good visualization takes time, dedication and patience. Failure of doing so will always result in a confusing one, which is neither efficient nor user-friendly.

Later in the context, you will see the plot we chose as a bad example, and the methods we used to make it better. They would help you better understand the concepts above.

## The Plot

### Introduction



[Figure 1] The original plot we chose.

The plot came from [China Daily](#), and is published in 2017, representing the research on Chinese short video market. We only cropped a part of it since we want quality over quantity.

## Background Stories

The reason why we took this plot is that it represents the booming of short video market in China, which we have witnessed in the past few years. Also, as short videos overwhelmingly spread across the country, more and more people are getting addicted to them, enjoying the ephemeral stimulation sent from each video.

This phenomenon caught our attention, and led to wonders about its rapid growth and when it started booming. The plot happened to be describing things about short videos, and thus got our attention.

## Analyses

This is a *Pareto Chart* with lots of decorations, mainly consisting of a bar chart, a line chart. The graph is presented in a double-rectangular coordinate system, with the left vertical coordinate representing frequency, the right vertical coordinate representing frequency, and the horizontal coordinate representing time.

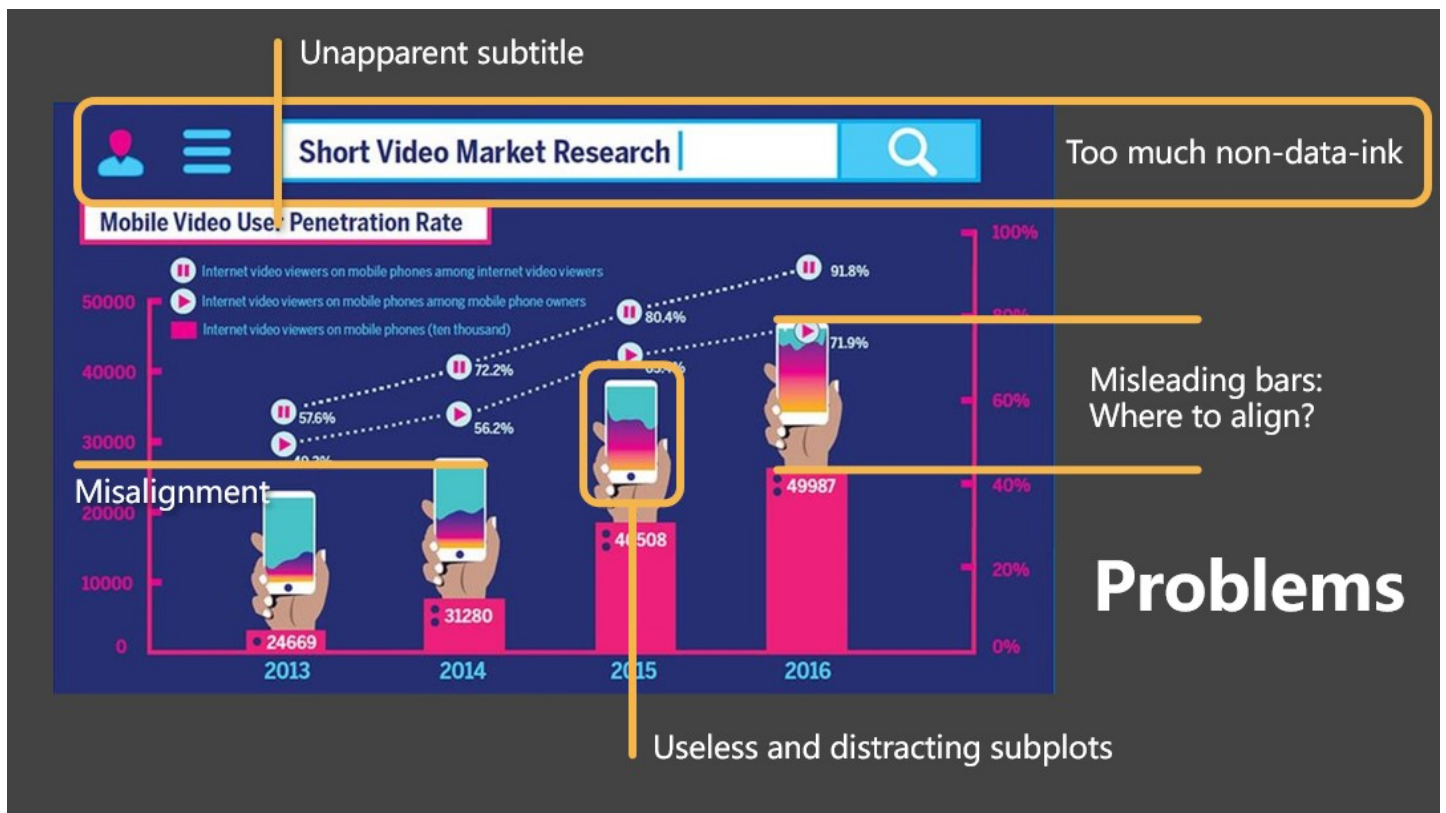
There are three variables in the graph and their legends, showing the trend of short video development in China from 2013 to 2016. The bar chart corresponds to the number of people who watched short videos in different years (ten thousand per unit) in the left y-axis, and the line charts correspond to the right y-axis, representing the proportion of people who watched short videos in different ways.

## The Problem

Here are the problems we found in this plot.

- **Unapparent subtitle:** The plot has a main title and a subtitle, and the main title is making the other one **unapparent**, which would cause a little confusion to the readers.

- **Unnecessary elements:** There are lots of unnecessary elements (a.k.a. non-data-ink, which does not transfer any useful information but takes much space), which prevent readers from retrieving useful information quickly.
- **Misalignment:** The bars are misleading. Readers will not be able to know where it ends at first sight (please refer to *Figure 2* for better understanding). It is hard to distinguish until they read the ticks on y-axis and the labels on the bar.
- **Misleading bars:** The bars are not correctly aligned. Take the bar at 2014 for example, the label reads 31,280, but the actual value it represents was **lower** than 30,000. This is a critical problem and should be avoided.
- **Useless and distracting subplots:** The subplots in the phones do not convey any useful information and have caused a lot of **distraction**.



[Figure 2] Problems found in the plot.

## Preparations

In the previous project, we used a *WeChat* group for communication and progress tracking. But there were some problems.

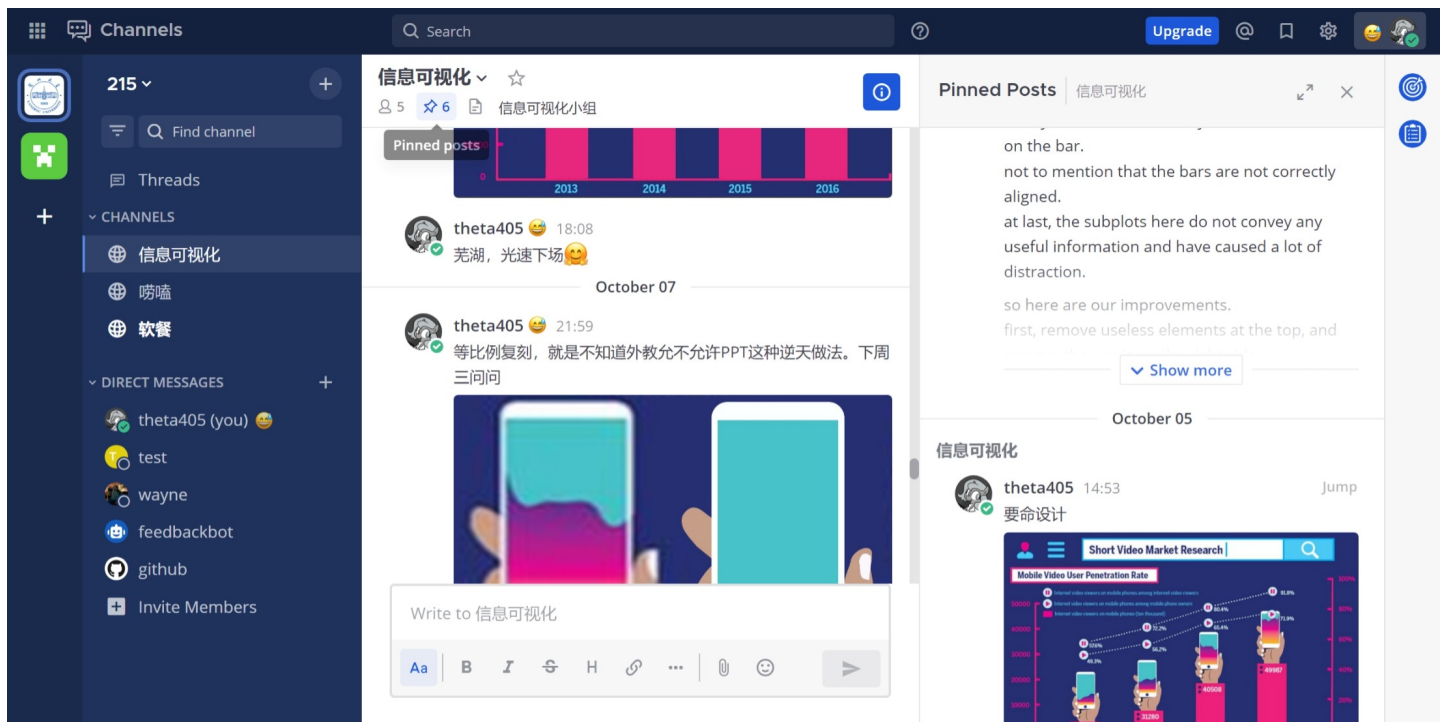
For example, chances are that your messages would **not** be synced to other team members' terminals, and thus resulting in a total **chaos** that throttles the development process.

So, for **communication**, based on lessons learned the hard way, we have switched to a self-hosted *Mattermost* server, which is an open-source alternative of *Slack* and meets our requirements quite well.

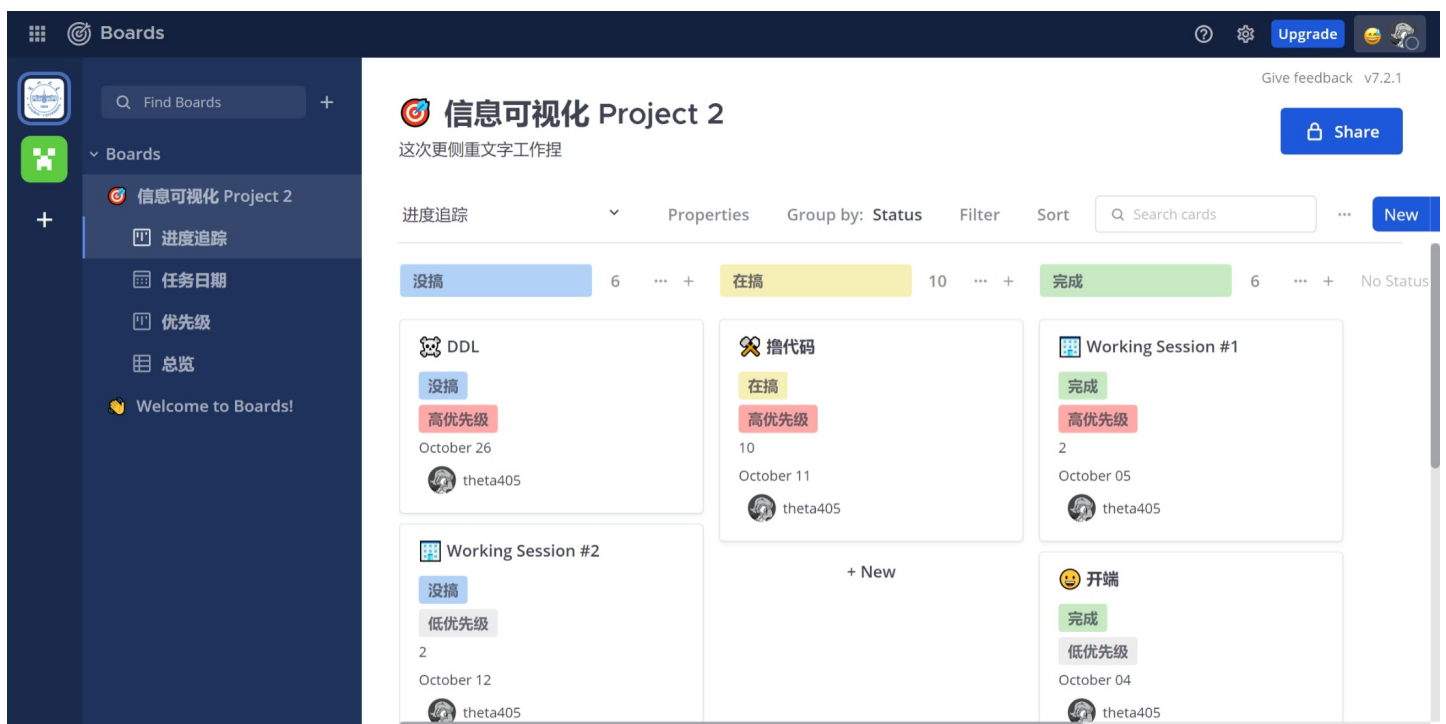
For the **tool** we used, after knowing that *Python* is not the only choice, we have decided to use *PowerPoint* instead, since it is easier to replicate complex plots.

For **file syncing**, we chose *Office 365* and *OneDrive* instead of *Git* and *GitHub*, due to the limited support the latter two have for *PowerPoint* files.

Here are some screenshots taken when the project started.



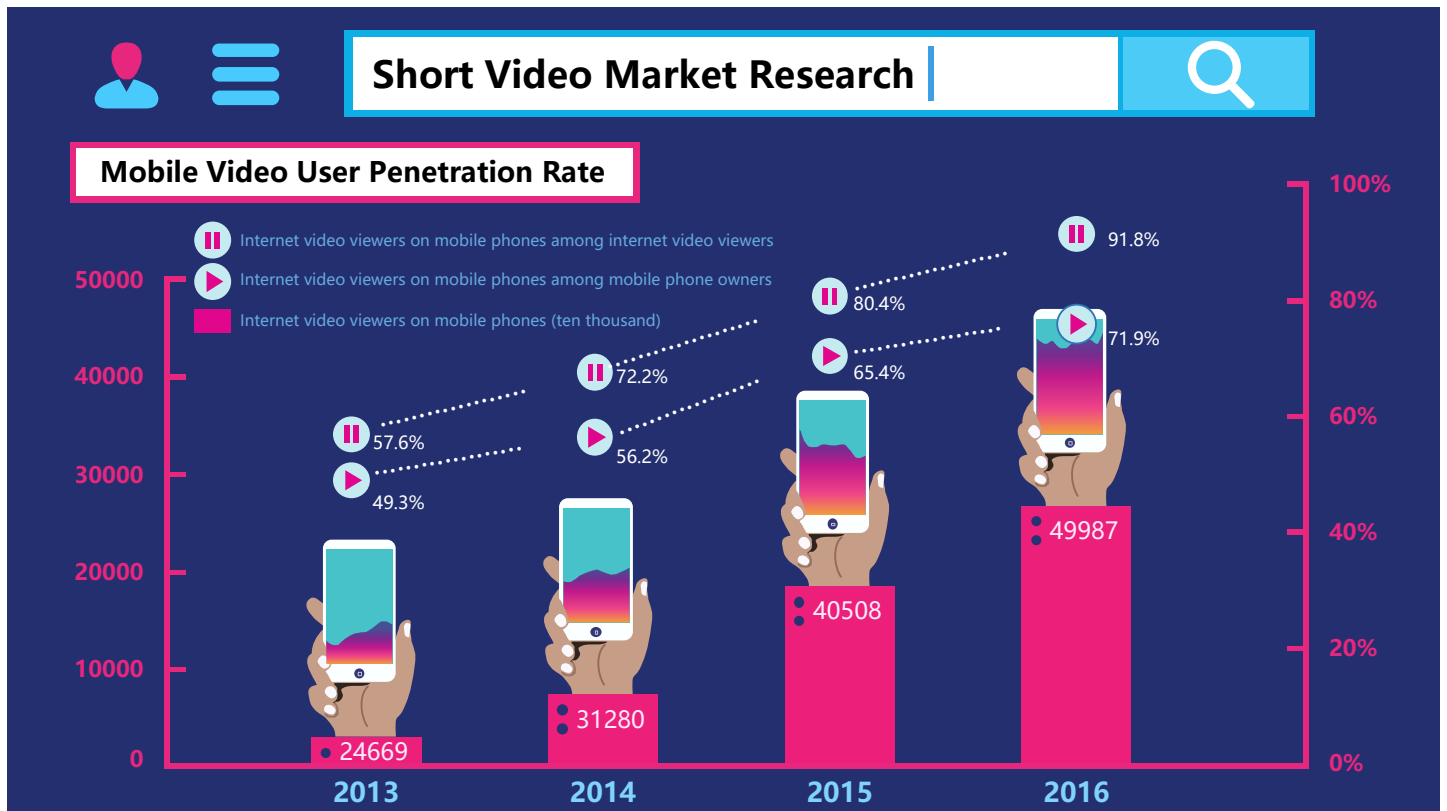
[Figure 3] Team members using mattermost to share their ideas.



[Figure 4] The board used to track progress.

## Replication

The replication is not very hard, but it did takes lots of patience due to tedious works like tracing free shapes in the plot using *free forms*. Eventually, we managed to finish it, and the result is quite satisfying.



[Figure 5] The replica we made using PowerPoint.

⚠ It seems that PowerPoint does not work with vectors very well ⚠

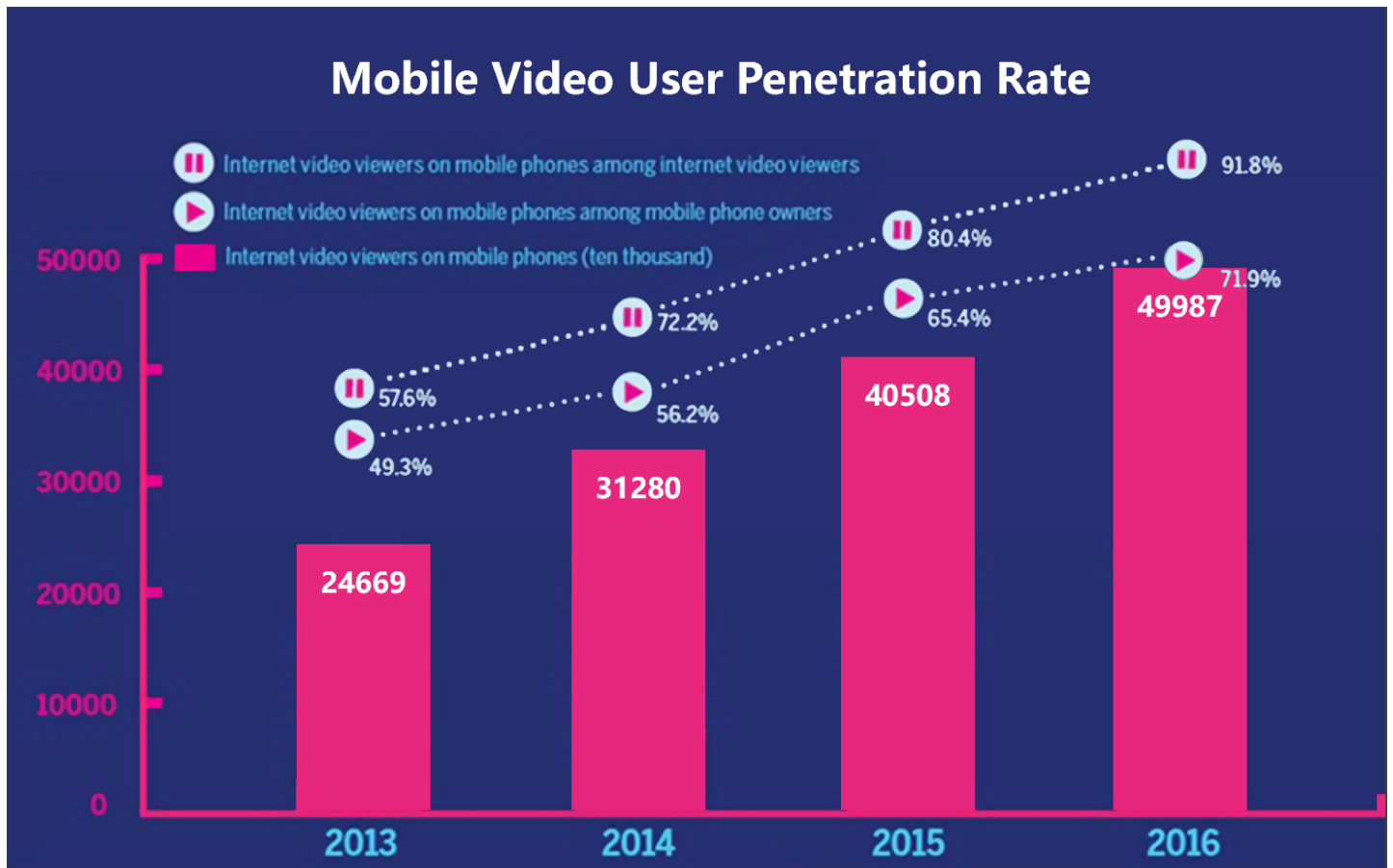
The dotted lines here are incorrectly displayed, and sadly there are no quick fixes.

We can assure that this is 100% correct in the slides and the exported pictures attached in "result pictures", and we are sorry for this error!

## Improvement

### First Version: Quick Fixes

Plot chosen, we began discussing what could be done to improve them. After some brainstorming, the first version was quickly made for our first working session and acted as a prototype. It is mainly made by adding, removing and moving elements from the original plot.



[Figure 6] The first improved version.

Soon, one of our member listed some problems that occurred in this version. The removal of the right y-axis made the line chart more confusing, because readers would not know why it was there. Also, the color and contrast is not ideal, parts of the plot are hard to read.

## Changes

Hence, we embarked on a journey of editing and revisions. All the changes we made along with their explanations (the italic texts) are listed below.

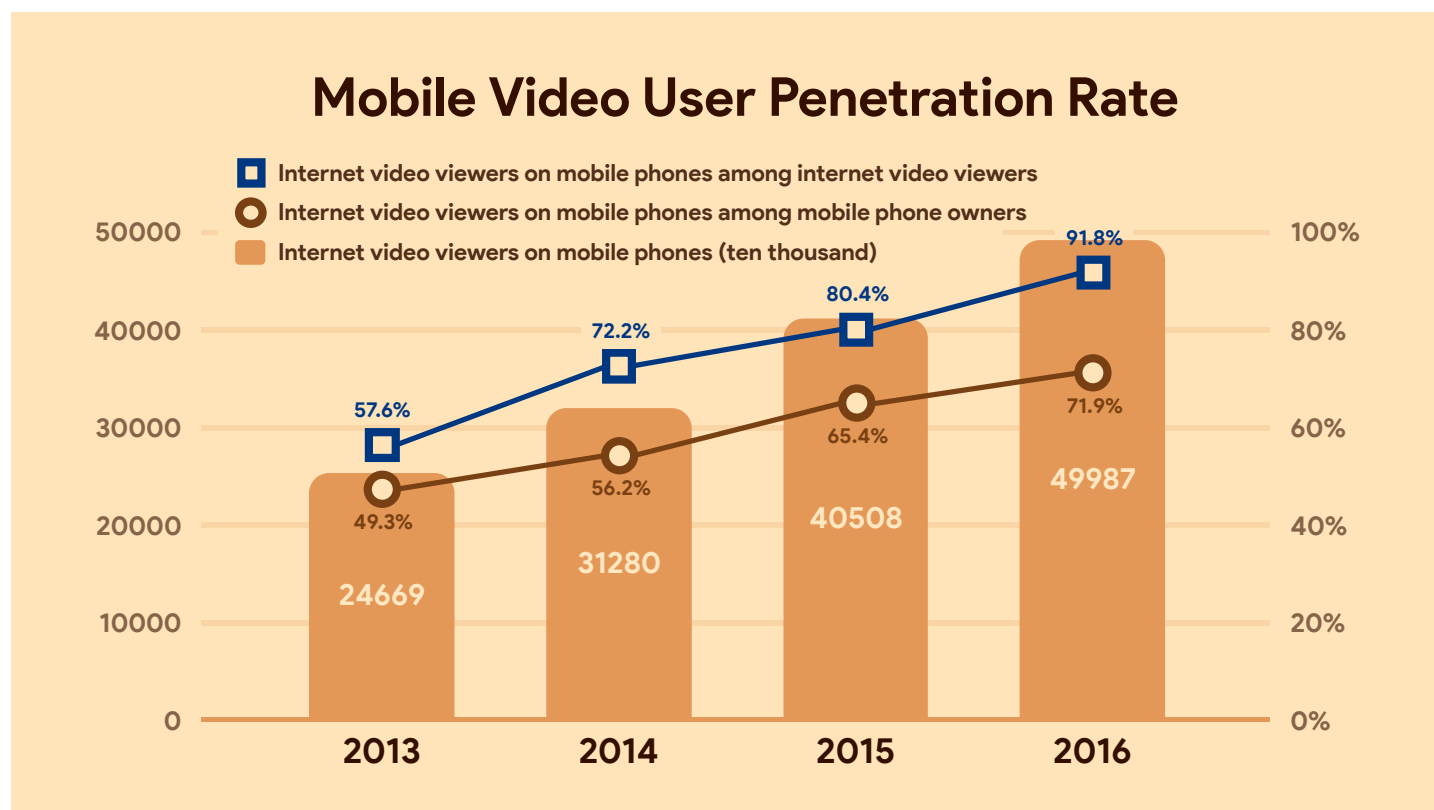
- **Removing unnecessary elements:** We removed all the unnecessary elements such as the title and the icons beside it. *This is done to make the plot look nice and clear, without any redundancy..*
- **Removing y-axes on both sides:** We removed the axes and added some horizontal grids. *The removal would not cause any problems since the grids would be the substitute, it also make the plot look clearer.*
- **Lower the right y-axis:** *It is lowered to make its ticks align with those on the left y-axis. The original plot changed the heights to compromise with its layout, while this is not a problem on the improved one.*
- **Changing width of the bars:** We changed the width of the bars and the gap between them. *Gaps too wide make the plot look empty, and it is advised that the width of gaps be half of the bars'.*
- **Changing the palette:** We changed the palette of the plot according to some presets on the Internet. The colors are carefully chosen so that they are all colorblind-friendly. *The palette of the original plot is not very satisfying, and we decided to replace it with some warm colors.*
- **Changing the font:** *The new font, Product Sans from Google makes the plot look nicer.*
- **Changing markers:** *The markers on the original plot has some unnecessary elements that would cause confusion. The new ones are different in both shapes and colors to make them more distinguishable.*



All of these changes can be applied to other plots if they share **the same** problems with our original one.

## The Final Version: Rejuvenation

After the whole working progress, this is the final version of the improved plot.



[Figure 7] The final version, made using *PowerPoint*.

Looking at this final plot, it complies with the principles we learned from the class. It might not be perfect (for example, the labels on the bars would be better if it is put at the top of the bar), but we tried best polishing as much details as possible.

## Conclusion

Long story short. Though we still encountered many problems doing this project, we have managed to conquer them one by one. This project taught us lots of valuable lessons that we would always benefit from in the future.