

Pulpulak data report draft

Hrach Khachatryan Lilit Ivanyan Adam Simonyan

2023-11-15

In this particular section we will answer the following quantitative questions:

1. What is the portion of pulpulaks that have a functional valve vs ones that have continuously flowing water? Produce graph(s) to highlight your result.
2. Taking into account the flow rates from continuously flowing pulpulaks in your dataset, how much water is lost per day in Yerevan? How much per year? What are those numbers equivalent to?
3.
 - a) Describe the distribution of pulpulaks between the different districts. Highlight if there are significant spatial discrepancies in their distribution density. Analyze the potential reasons that lead to these differences.
 - b) Develop a methodology to identify where it is more likely to find pulpulaks. Try to find correlations related to some of these factors:
 - a. Road types
 - b. Architecture
 - c. Socioeconomic realities and institutions
 - d. Proximity to certain institutions and entities
 - e. Demographics
 - f. Geography

Do your conclusions somehow explain the spatial distribution answered above?

4. Are there significant differences in flow rates of continuously flowing pulpulaks between different districts? Hypothesize why there may or may not be a difference depending on your answers.
5. What is the portion of pulpulaks that have a functional valve vs ones that have continuously flowing water? Produce graph(s) to highlight your result.

1. What is the portion of pulpulaks that have a functional valve vs ones that have continuously flowing water? Produce graph(s) to highlight your result.

After importing the data and filtering it we will get only the information for pulpulaks located in Kentron district of Yerevan since that is the only district we are working on at the moment. The R programming language and R studio were used to complete the data analysis and all the codes that were run for the analysis are available in the ANEX section of the report.

Getting the absolute numbers of:

1. Pulpulaks with valve
2. Pulpulaks with no valve
3. Pulpulaks with a broken valve
4. Total number of pulpulaks

Here are the categories that pulpulaks have been split into based on the existence of valve:

```
## [1] "no valve"                "with a broken valve"
## [3] "with valve"             "1. no valve, 2. closed head"
## [5] "1. with valve, 2. no valve" "3 with valve, 1 with no valve"
```

The categories with several values in them were removed for data consistency. These represent just a couple of pulpulaks and would not make much of a difference if we get them removed.

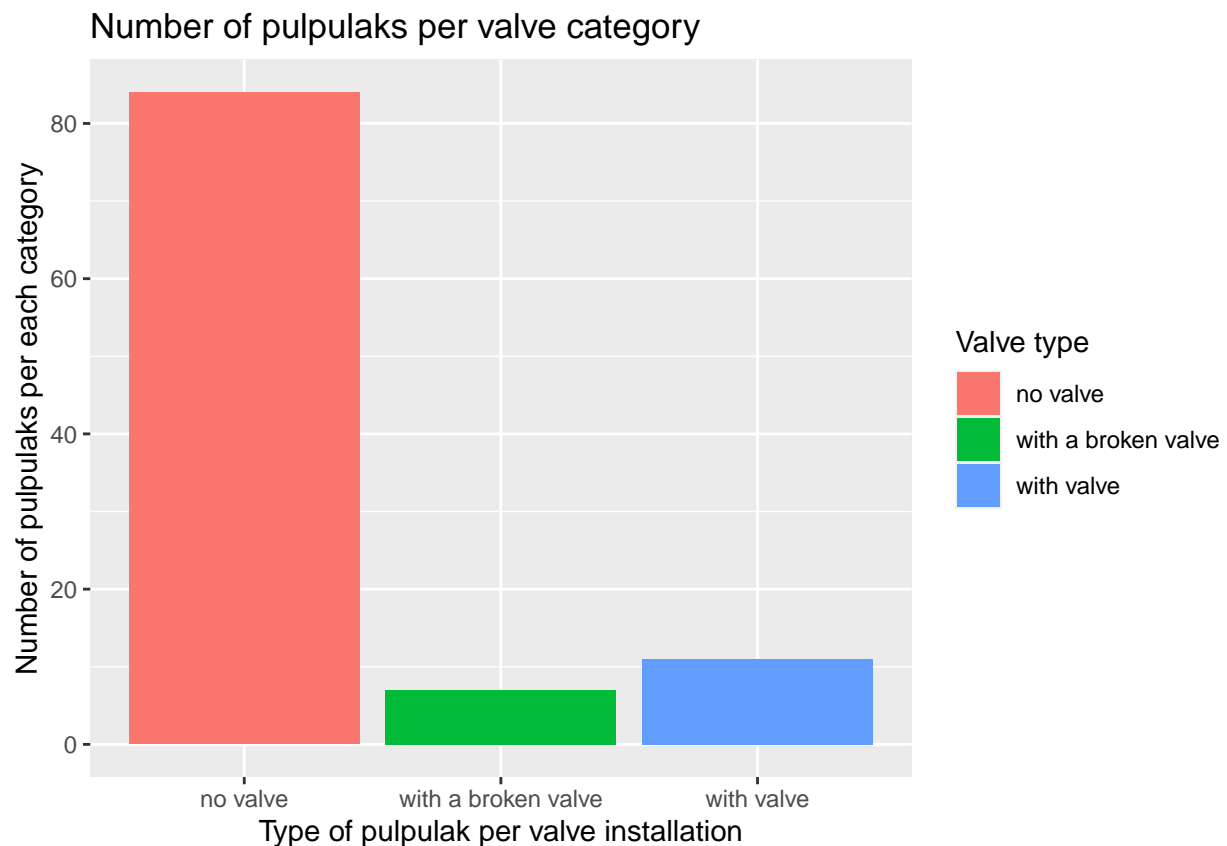
We need to get the absolute numbers per category:

Number of pulpulaks with a valve: 11

Number of pulpulaks with no valve: 84

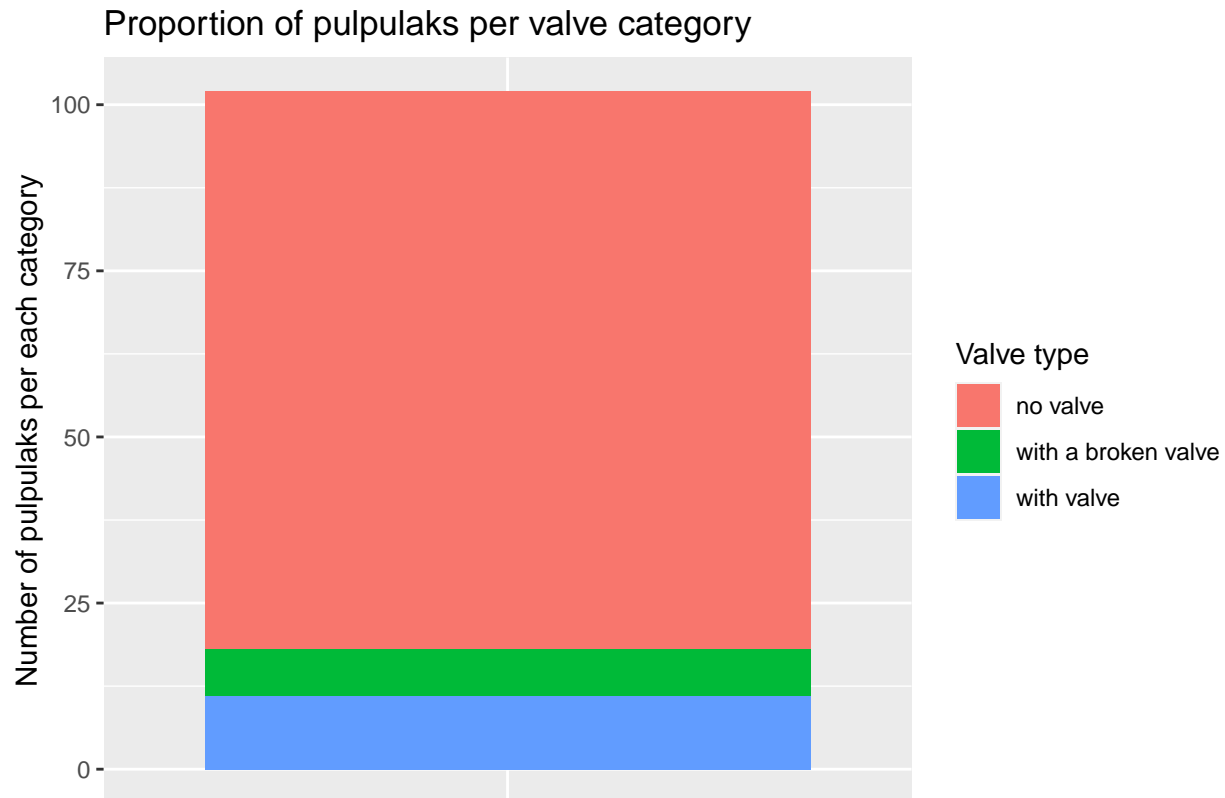
Number of pulpulaks with broken valve: 7

Now, let's draw some graphs to look at the proportion and distribution per categories. First we will draw a barplot:



Now let's look at the barplot in another form for proportion comparison:

```
ggplot(data = filtered_data, aes(x = factor(1), fill = Type)) + geom_bar() +
  labs(title = "Proportion of pulpulaks per valve category",
       x = "",
       y = "Number of pulpulaks per each category",
       fill = "Valve type") +
  theme(axis.text.x = element_blank(),
        axis.ticks.x = element_blank())
```



2. Taking into account the flow rates from continuously flowing pulpulaks in your dataset, how much water is lost per day in Yerevan? How much per year? What are those numbers equivalent to?

Let's get to the amount of water lost in Yerevan per day and per year and compare it with other forms of water usage. We will obviously exclude the pulpulaks with valves since they are not wasting water.

Exporting the data into excel for further manual refining:

Importing the manually refined data:

Let's get the numbers for water wastage:

Here is the total amount of water wasted in Yerevan:

1. Per minute: 353.31
2. Per day : 5.29965×10^5
3. Per year: 1.9343722×10^8

The wasted water in Yerevan per day is roughly equivalent to the per capita water usage of about 2,902,824 people in New York City per day.

The wasted water in Yerevan per day is roughly equivalent to the per capita water usage of about 5,186,218 people in Tokyo per day.

The wasted water in Yerevan per day is equivalent to about 0.0015% of the daily water usage in the entire United States.

The wasted water in Yerevan per day is equivalent to about 0.00003% of China's annual water usage per day.

The wasted water in Yerevan per year is roughly equivalent to the per capita water usage of about 1,059,178 people in New York City.

The wasted water in Yerevan per year is roughly equivalent to the per capita water usage of about 1,891,455 people in Tokyo.

United States: The total water usage in the United States is around 3,800 billion gallons per day (approximately 14,384,000,000,000 liters per day). The wasted water in Yerevan per year is equivalent to about 0.0014% of the daily water usage in the entire United States.

China: China's total water usage is around 554 billion cubic meters per year (approximately 554,000,000,000,000 liters per year). The wasted water in Yerevan per year is equivalent to about 0.00003% of China's annual water usage.