

Visualization for Navy Hearing Conservation Program (HCP)

Hung Do, Srinivas Havanur, Erika Siregar

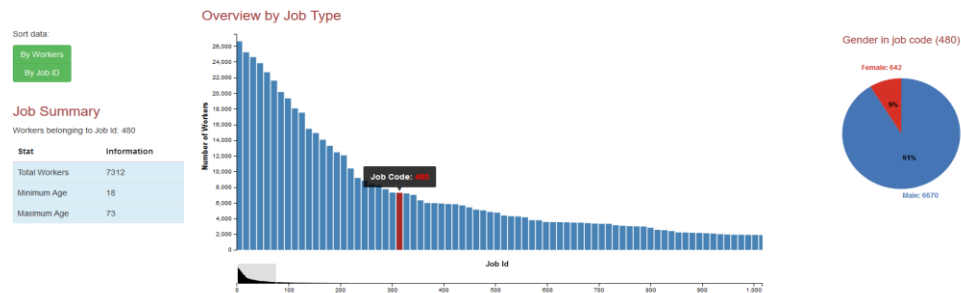


Figure 1. Overview graph of Navy Hearing Conservation Program.

Abstract- The aim of this paper is to track hearing level of workers in US navy over a period of time through Hearing Conservation Program (HCP). Hearing Conservation Program is an internal training conducted by the US Navy with a goal to detect and prevent a noise-induced hearing loss among the service members by analyzing their hearing levels over the years. The authors analyzed the data obtained from the Audiogram dataset to produce some interactive visualizations using D3.js to see hearing curves of workers over the years.

Index Terms- Hearing level, navy, audiogram, noise

1 INTRODUCTION

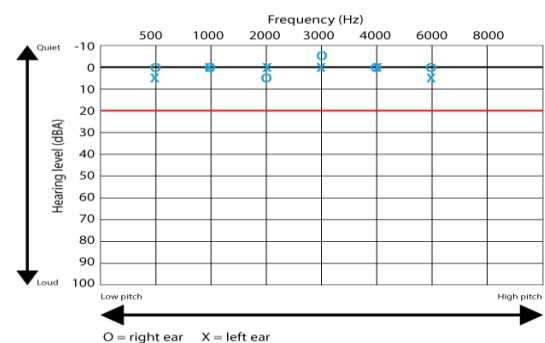
Hearing Conservation Program (HCP) conducted by US Navy aims to protect and prevent the hearing loss in service members. This program is based on the idea that being exposed to an extremely loud noise could cause someone instantly lose the hearing ability. However, people who are exposed to the noise for a very long term will have a five-year lag before the hearing loss could be noticed.

People who are tested in this program are workers who exposed to noise levels in the range 85-100 decibels (dBA). Their hearings are tested on regular basis. The test details are recorded in the audiogram so that there will be beep sound at different frequencies as and when the volume is increased. Figure 1 shows an audiogram for normal ears.

A person who is taking the test raises their hands as a response after hearing the beep and the frequency along with the volume are recorded. Hearing loss is indicated by a higher dBA number, which is recorded by an "O" (right ear) or "X" (left ear). Figure 2 shows an example of audiogram for noise-induced hearing loss. Due to limitations in the data collection methods and the characteristics of the Navy worker population, it is difficult to learn more about the data by using traditional data analysis. Workers often do not stay in the same job for many years. For example, if a worker changes his job from the flight deck of an aircraft carrier to

the laboratory and then has his hearing measured, it might appear that lack of noise exposure was causing his hearing loss. Therefore, a visualization is needed to help better understanding the audiogram data.

In this paper, the authors analyzed a dataset of audiograms having more than 700,000 records covering 20 years of the program till the year 2012 and produced several interactive visualizations using D3.js for the user to see the hearing curves of workers over the years. Hearing curves are visualized by JobID and the WorkerID. For hearing curves by JobID, overview-detail charts are provided to identify the average hearing levels for different age ranges. In addition to these, the heat map is provided to see the frequency of the workers belonging to that JobID. Another overview-detail chart is designed for hearing curves by WorkerID. This visualization allows the user to explore hearing curves for individual workers at different age along with their total hearing curves. This will be discussed in more detail in further sections.



- Hung Do is with Department of Computer Science at Old Dominion University, Email: hdo@cs.odu.edu
- Srinivas Havanur is with Department of Computer Science at Old Dominion University, Email: shavanur@cs.odu.edu
- Erika Siregar is with Department of Computer Science at Old Dominion University, Email: esiregar@cs.odu.edu

Figure 1. Normal audiogram

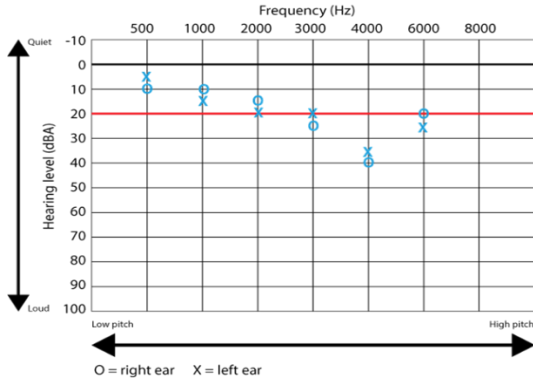


Figure 2. Noise-induced hearing loss audiogram

This paper seeks to help answering some of the questions as follows:

- Identify various shapes of the hearing curves for every worker exposed to different frequencies.
- Know the individual's hearing curves over the period of time.
- Identify the shapes of the hearing curves of age group or subgroup.

2 DATASET

The original dataset is CSV format which is obtained from the audiogram having 734,327 records. It consists of many attributes, some of which are considered more interesting for producing visualization, such as Gender, JobId, WorkerID, Age, and Hearing Levels at different frequencies.

To clean the data, we use R to find the items that have null values and then remove them. Because these null values would affect our derived data. In this case average hearing levels and age of a worker are derived attributes which are also computed using R. To reduce the time for data access, filtering of data is provided using MySQL and PHP scripts to load the visualization with greater response time.

3 OVERVIEW OF THE VISUALIZATION

The visualizations are built using d3.js and jQuery. In addition, MySQL and PHP are used to store the dataset. Also, a website template is developed by using HTML, CSS, and Bootstrap. All these have been listed in the references section.

The complete visualization project can be seen at <http://www.cs.odu.edu/~hdo/InfoVis/navy/final-project/index.html> and the live demo is available at <https://youtu.be/twAitZyBllo>.

3.1 Overview by JobId

An overview bar chart showing 1018 JobID in the program allows the user to get general information from that JobID (job type) including male and female statistics. Also, additional information for each JobId is provided by

hovering the mouse on one of the job types to get details in a tooltip and summary table.

This visualization has several method of interactions. A slider with context and focus allows the user to select an area of interest of all job types. Mouse actions (click and hover) provide information in detail for that job type. In addition, options are provided to order the bins by JobId or number of workers to help the users to see the interesting job details.

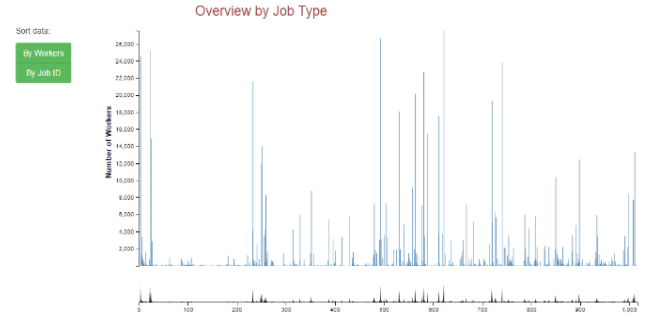


Figure 3. Overview number of workers in one JobId.

3.2 Heatmap for a particular job

This graph shows the distribution of workers by age and total hearing belonging to a particular job type. It provides an ability to explore patterns of workers in one job type. Also, it presents information about the total hearing range that is common for different workers by age. The tooltip for details is provided by hovering the mouse on each point on the heatmap.

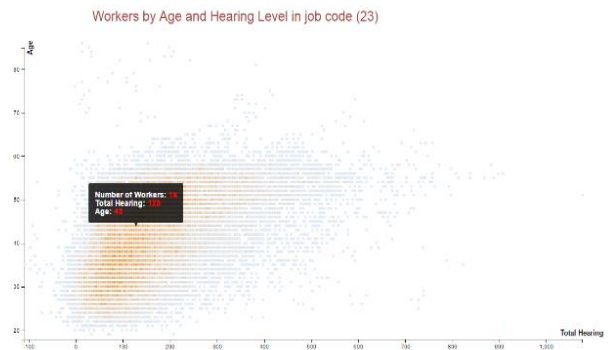


Figure 4. Heatmap to show the distribution of workers by Age and Total Hearing in each JobId.

3.3 Average hearing level by age range

The following line graph shows average hearing level curves for particular job type (JobID: 232) by different age or age group selection. Hearing curves includes different frequencies exposed to left and right ears which are encoded with different colors as shown. The chart area is divided into six regions represented the hearing levels: normal (-10 dB - 25 dB), mild (26 dB - 40 dB), moderate (41 dB - 55 dB), moderate-severe (56 dB - 70 dB), severe (71 dB - 90 dB), profound (91 dB - 110 dB). The different regions of hearing level are encoded using the saturation of blue.

Using a slider, the user can select an age range and see the average hearing curves of that range. This graph takes JobID as input from the overview graph through the clickable event.

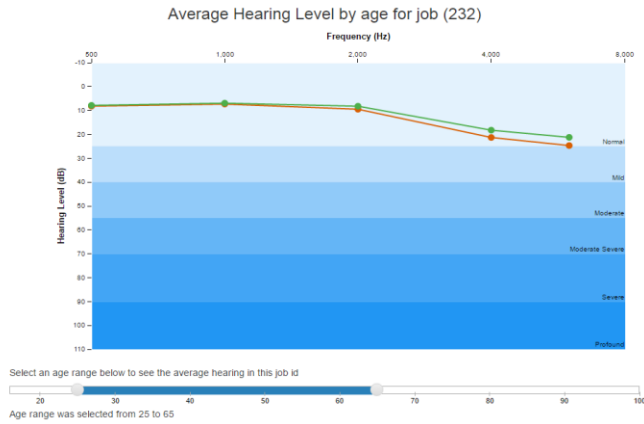


Figure 5. Average hearing level by age range selection (slider).

3.4 Overview by worker

This graph represents the overview of individual workers. The dropdown filter is added to see the overview graph by birth year, by this way, the large dataset having more 700,000 records can be presented faster in the visualization. The user can select interested range of WorkerIDs from the graph through an area shaped slider. In addition, hovering the mouse on individual bar plot gives summary details for that WorkerID.

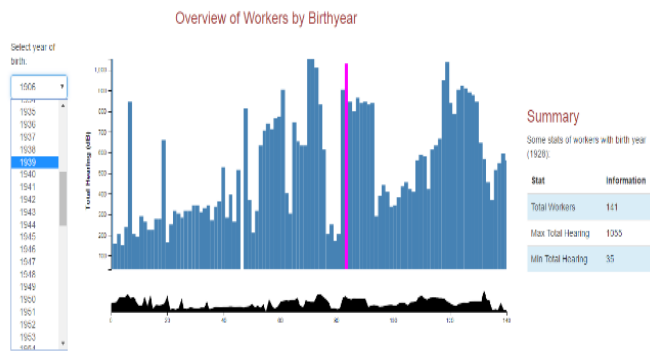


Figure 6. Overview of workers by birth year.

3.5 Hearing levels in detail of each worker

The graph in figure 6 takes a WorkerID in figure 5 as an input through the clickable event to show hearing curves of a particular worker over the years in the program. A slider is provided to see the hearing curves of the worker for every test date over the period of years as shown. In addition, hovering the mouse over individual points will bring up a tooltip showing the details of either left or right ear along with their plotted x and y values.

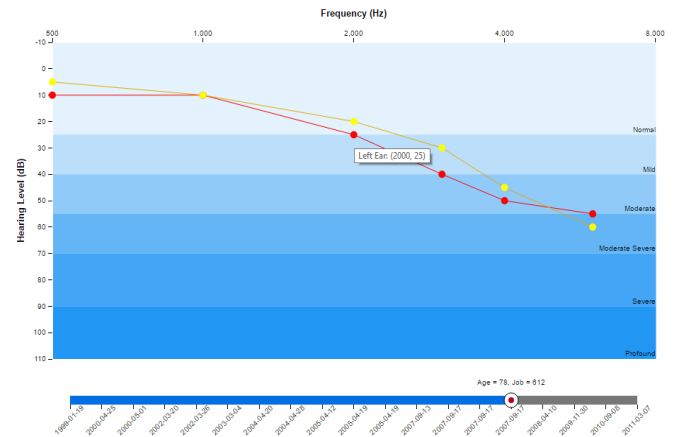


Figure 7. A graph of hearing levels for a particular worker.

Moreover, a graph displaying total hearings over the time of a worker is provided in the graph in figure 7. If a worker changes his job, this chart will show JobID with different colors. Also, a tooltip is provided to see the details of the JobID.

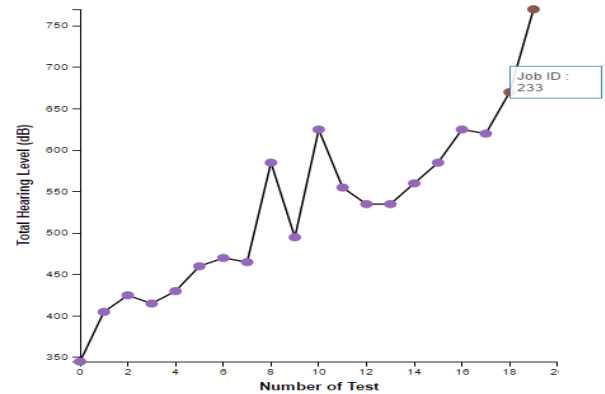


Figure 8. Total hearing levels of one worker over the time.

4 WHAT-WHY-HOW FRAMEWORK

The what-why-how framework described by Munzer acts as a basic guideline and framework for visualization any form of data. We will discuss how this framework was effectively applied in our project and a short summary of this can be found in Table 1.

4.1 What: Data

The original dataset is a multiple-attributes table in a CSV format. This dataset consists of 53 attributes (quantitative and qualitative) and 734,327 records with one key attribute, which is WorkerID (sixplace_id). Because of its size, this dataset was, then, exported into MySQL database which is running on ODU Server.

4.2 What: Derived

Two derived attributes are computed:

1. Average hearing level for each selected age range.
2. Age of each worker based on their year of birth and their test date.

System	Navy Hearing Conservation Visualization
What: Data	Table: - Multiple quantitative attributes - Multiple qualitative attributes - One key attribute (WorkerID)
What: Derived	- Average hearing level - Worker's age
Why: Tasks	- Present the overview of the dataset - Locate a particular worker - Compare the hearing level of people from different age - Identify patterns and find correlation between hearing level and age - Compare the hearing levels of different jobs. - Identify hearing level pattern for each worker - Summarize data distribution
How: Encode	- Bar chart - Pie chart - Multi-line chart - Single line chart with circle's color representing the job type - Scatterplot with heatmap - Area chart for overview's slider
How: Manipulate	- Select interesting bin - Select certain number of bins to focus on - Zoom and pan the bar chart - Reorder (by workers, by JobID)
How: Facet	- Overview - Detail - Popup detail view - Juxtapose multiple views - Share navigation: synchronize
How: Reduce	Filtering (by year of birth)

Table 1: What-Why-How framework

4.3 Why: Abstract Tasks

There are three main actions that users can do with this visualization, which are *analyze*, *search*, and *query*. In *analyze*, users can present the overview of data so that they will be able to choose interesting items from the data for further analysis. In *search*, users can locate the workers of interest using the zooming and panning features. In *query*, users can identify the pattern and find a correlation between hearing level and age, and identify hearing level pattern for each worker. Users could also compare the hearing level of people from different ages and the hearing level of different jobs as well as summarize the dataset (total number of workers by gender, maximum hearing level, minimum hearing level, etc.).

4.4 How: Encode

Our visualizations are grouped into three main categories:

1. Job overview chart and workers overview chart. These charts are encoded using line marks and aligned vertical position to express value, separate key attribute with horizontal position. There are three colors being used, which are blue for the non-selected bin, purple for selected bin, and dark orange for hovered bin. The job overview chart is accompanied with a pie chart showing the number and percentage of workers for a particular job id. This pie chart is encoded using area marks with angle channel and radial layout.
2. Detail chart for workers by age and hearing level in a particular job. This chart is a combination of scatterplot and heatmap, which is encoded using horizontal and vertical spatial position, point marks, and diverging colormap.
3. Detail chart for the hearing level of individual worker. This chart is a line chart encoded using point marks that are connected with line marks. The lines are distinguished using two colors: red for left hearing and yellow for right hearing. The background of this chart is colored based on the type of hearing level (normal, mild, moderate, moderate severe, severe, and profound).

The chart for individual hearing level is complemented with another line chart showing total hearing level. The point marks are colored according to the job type that is represented.

4.5 How: Manipulate, Facet, Reduce

To manipulate the dataset, users can select how many bins they want to see using the slider-area chart and then select a bin of interest by clicking that particular bin. Users could also see the bins in more detail by zooming and panning the chart and reorder the bins according to the number of workers or job type.

For facet, we use overview and detail; juxtapose multiple views and share navigation: the right-left hearing level and total hearing level), overview bar chart and pie chart; and popup detail view when selecting or hovering over a certain area in the chart. The user can filter the dataset of interest by selecting a certain year of birth from the drop-down list.

5. RELATED WORK

Some interesting visualizations have been produced by Lulwah Alkwai, a Ph.D. student from ODU. These visualizations include hearing level by job type, and a scatterplot of age and total hearing for workers.

These works and some documentations related to this were really helpful to understand the dataset of HCP. In addition, some existing technical terms and representations have been reused in our visualizations. These includes hearing levels, noise exposure, profound level etc.

6. CONCLUSION

Interactive visualizations are produced by making use of audiogram dataset to analyze the hearing curves to among the workers belonging to different JobID. Also, it could be used to analyze the average hearing curves of workers of different JobID belonging to different age group. We observed how these visualizations helped to identify the detrimental factors of hearing over the period of years. Moreover, the visualizations are not only providing aggregate information of the dataset, but also individual information for a particular worker.

These visualizations were designed using D3.js technologies and MySQL database is used to filter the workers by date of birth to produce the hearing curves for different age groups. Various interaction idioms have been used to give user-friendly experience. For our future work, visualization using some more interesting aspects of workers will be produced to identify and solve this issue of hearing problem that occurs over the time. Also, performance improvement of an algorithm will be made to give better user experience.

7. FINAL THOUGHTS

Creating the visualizations for the HCP was not an easy task. We faced many challenges that made us learned new knowledge and sharpened our skills in using d3. First, we struggled to understand the concepts that are used in the HCP. Second, we struggled to find the ideas about how to “translate” the requirement from the user into an easy-to-understand visualization, which could answer their questions. Third, we had to deal with a very large dataset, which affected the performance of our visualization. We worked hard to figure out how to handle all of the workers records without causing any “loading time” issue.

The dataset used for the implementation is huge having more than 700,000 records which might have caused the bottleneck for the visualizations that are produced. Some good approaches are adapted to deal with this situation

which includes the use of range sliders to produce visualization using Focus+Context via brushing.

Also, some performance overhead is improved by using MySQL to store the dataset and built an API system (PHP and JSON) that allows visualizations accessing dataset via these APIs. By using this technique, we avoid the delay when interacting with visualizations.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. John Muller, Occupational and Environmental Medicine Navy and Marine Corps Public Health Center, Portsmouth, VA for providing some documentations and dataset to analyze their requirement to represent the hearing curves patterns for different workers in the navy to identify the problem of hearing loss over the period of time. The authors also would like to thank Dr. Weigle, Associate Professor, Computer Science, Old Dominion University for providing us great ideas on how to produce good visualizations and for guiding us in the right direction. Finally, the authors wish to thank Lulwah Alkwai for providing the initial visualization of the hearing level and assisting us in understanding the concept and problem in the Navy Hearing Conservation Program.

REFERENCES:

- [1] Bootstrap: <http://getbootstrap.com/>
- [2] Data-Driven Documents: <https://d3js.org/>
- [3] JQuery: <https://jquery.com/>
- [4] Lulwah, *Navy Hearing Conservation Program Visualizations* <http://ws-dl.blogspot.com/2014/06/2014-06-18-navy-hearing-conservation.html>
- [5] Mike Bostock, *Focus+Context via Brushing*, <https://bl.ocks.org/mbostock/1667367>
- [6] Navy Medical Department Hearing Conservation Program Procedures, <http://www.med.navy.mil/sites/nmcphc/Documents/policy-and-instruction/oem-navy-medical-department-hearing-conservation-program-procedures.pdf>
- [7] Siemens, Degrees of Hearing Loss, <https://usa.bestsoundtechnology.com/hearing-loss-and-tinnitus/understanding-hearing-loss/degrees-of-hearing-loss/>
- [8] Slider, <http://sujetsr.github.io/d3.slider/>
- [9] Pie chart, <http://bl.ocks.org/enjalot/1203641>

AUTHOR BIOGRAPHIES

Hung Do got his B.S from Hanoi University of Science and Technology in 2002 and has been working in the field for more than 10 years of experience in Oracle, MySQL, and PHP. Currently, he is pursuing a master's degree in Computer Science at Old Dominion University.

Srinivas Havanur is a Graduate Teaching Assistant and a Master student in the Department of Computer Science at Old Dominion University, Norfolk, VA. His area of interests are Big Data, Data Visualization, Web development and Mobile app development using Android. He is having 3 years of industrial experience in UNIX/C++.

Erika Siregar is a Master student in Computer Science at Old Dominion University. She got her bachelor degree in Computational Statistics from Institute of Statistics, Jakarta. She has experience working in national statistics office in Indonesia for three years before deciding to pursue a master's degree.