VISUALIZATION PROJECT

Sales and purchases visualization

Submission Date – 31.8.16
Lecturer – Peter Bak
Students – Dana Borinski 305201238 & Danielle Israel 200684512

THE DOMAIN:

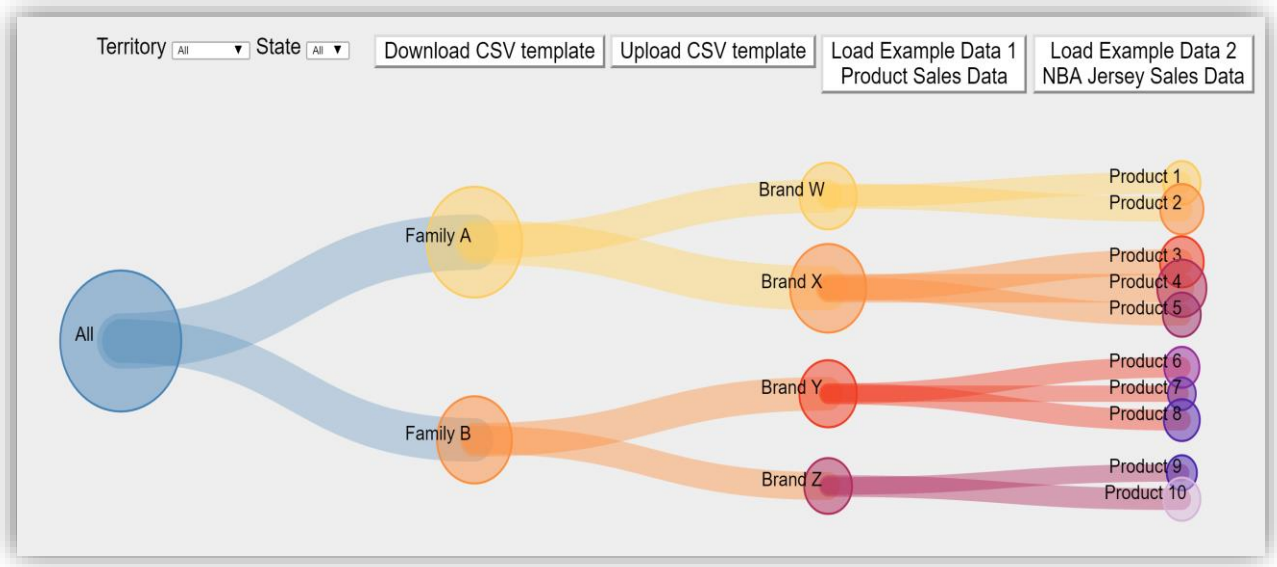
Our project deals with purchase-orders and sales-orders of a stand-alone business called 'שמואל'. This business sales office equipment and machinery.

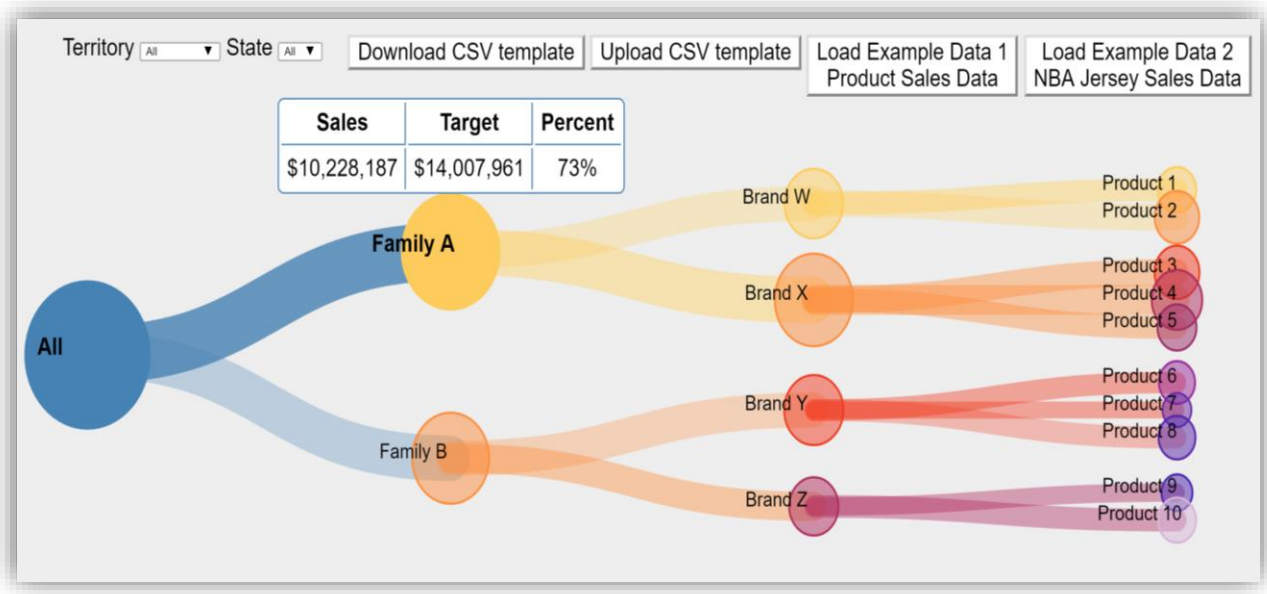
We've used a static data set of purchase-orders and sales-orders of the past year (since May 2015), while each order contains many rows of different items with different quantities.

For those orders we have chosen some of the most popular suppliers and customers of the business. Moreover, we've used statistics about sales and procurement of this business based on the owner's knowledge and experience to generate the data set.

SIMILAR VISUALIZATION'S DOMAIN EXAMPLES:

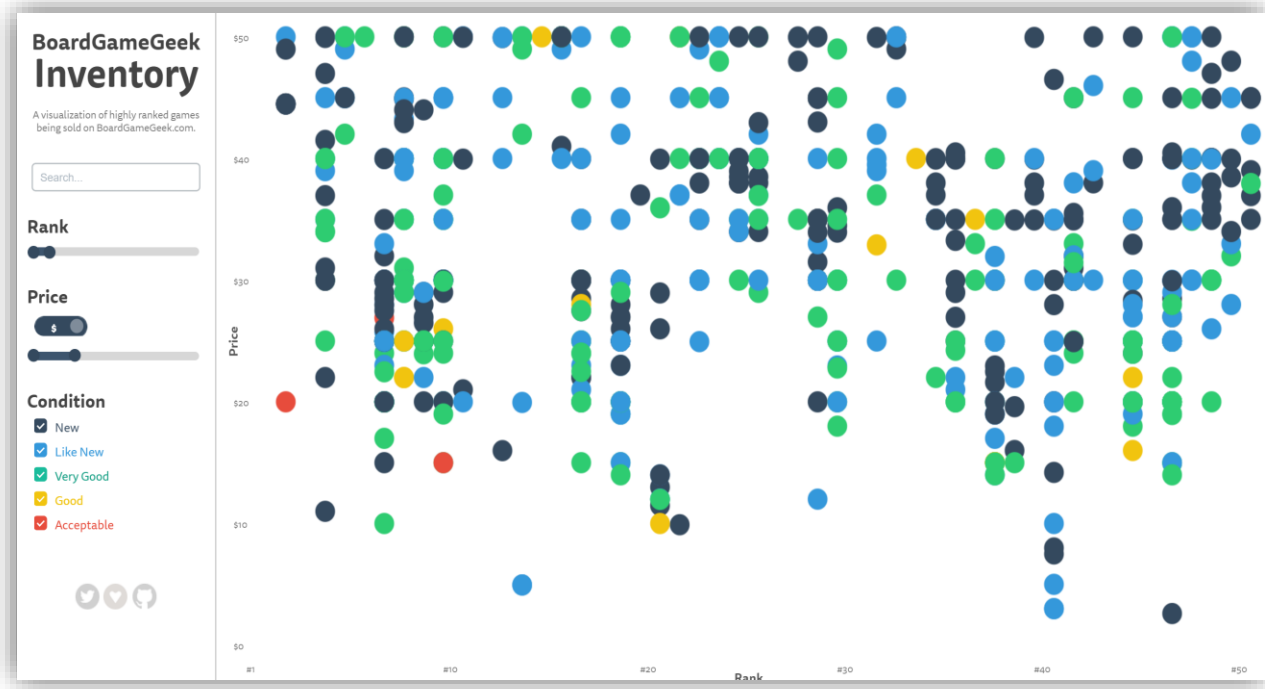
1. An example for sales-data visualization.





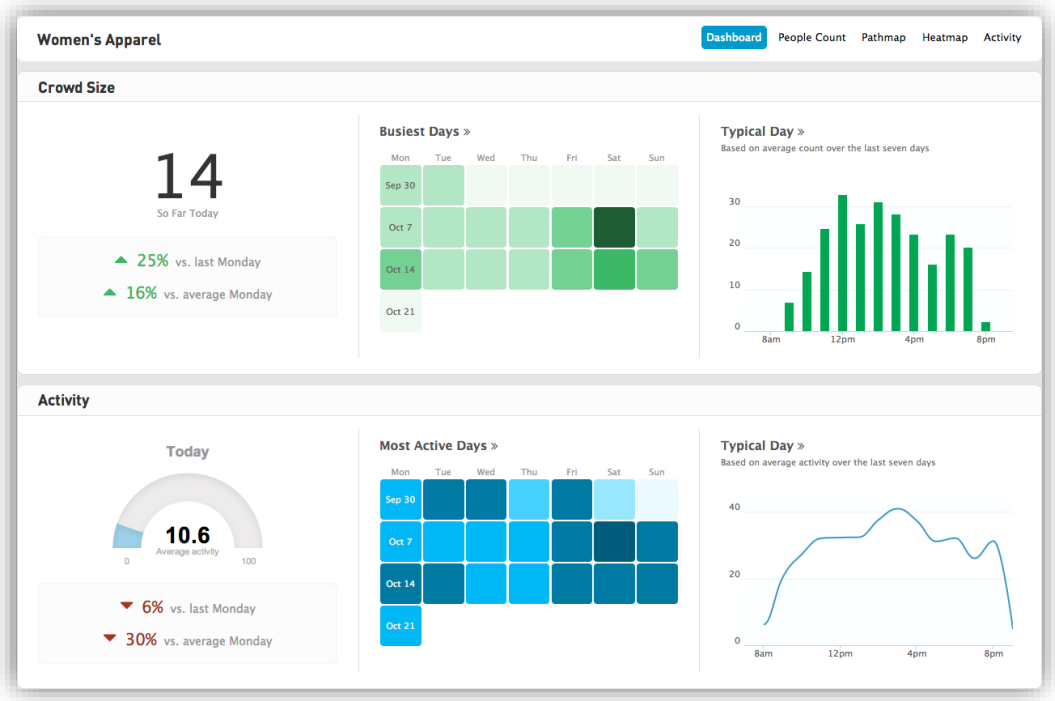
Reference: http://sales-data-visualization.herokuapp.com/

2. Board Game Geek Inventory - A visualization of highly ranked games being sold on BoardGameGeek.com.



Reference: http://bgg-inventory.com/

3. An example for Dashboard that includes different types of visualizations for customer's activity and movement at a store. To show 'most active days' or 'busiest days' they are using a calendar heat-map.



Reference: https://hbr.org/2014/05/how-data-visualization-answered-one-of-retails-most-vexing-questions?utm_campaign=May%25202014

THE DATA:

Data Attributes:

- Item number
- Quantity per item (purchased or sold)
- Price per item (purchased or sold)
- Total cost per item (purchased or sold)
- Total units per day
- Total cost per day
- Date (sales or purchases)
- Month and weekday

- Color (on a comparison of two items filter)
- The stroke (on a comparison of item or two items filter – the quantity divided by the total quantity)
- Size and hex (general view – quantity tabs)
- The stroke (on a comparison of item or two items filter – the item cost divided by the total cost)
- Size and hex (cost view – cost tabs)
- A specific position of a circle on the calendar heat-map and the tooltip for each day
- Labels for the calendar heat-map

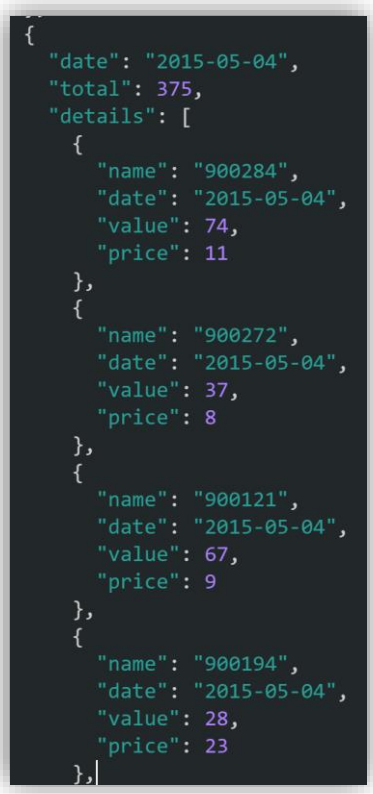
Data Attributes Types

Data attribute	Type
Item number	Ordinal data
Quantity per item (purchased or sold)	Quantitative data
Price per item (purchased or sold)	Quantitative data
Total cost per item (purchased or sold)	Quantitative data
Total units per day	Quantitative data
Total cost per day	Quantitative data
Date (sales or purchases)	Ordinal data (sequential) - Time data as an event, without duration
Month and weekday	Ordinal data (calendar) months- left to right – the columns, weekdays – top down – the rows. There is an option to hover on a month or a weekday and see only this column or row (the others become almost transparent)

Descriptive statistics:

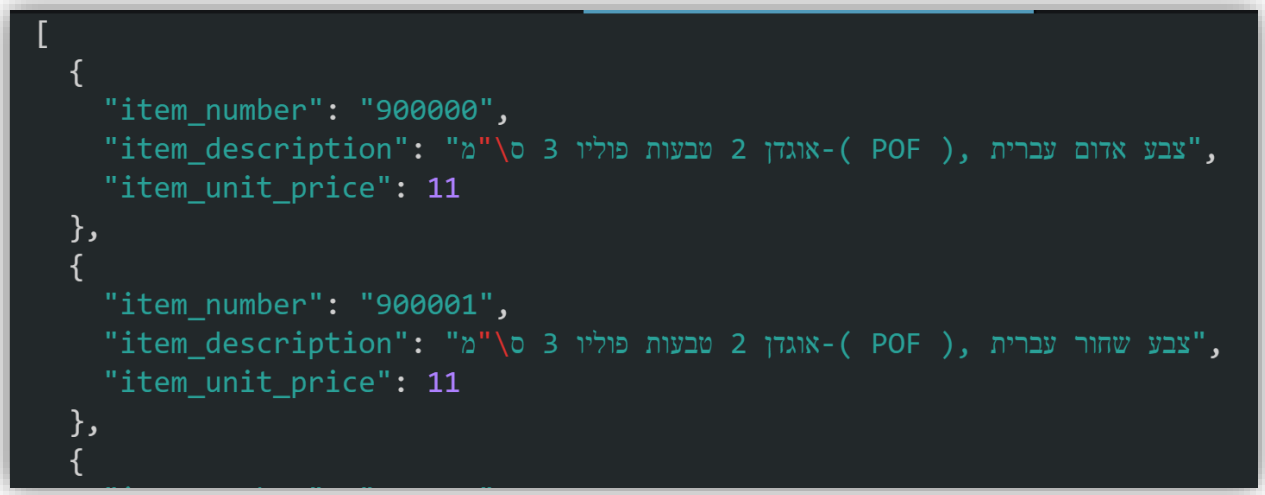
We used data of sales orders and purchase orders from 'י.שמואל' business. We organized the data we needed in a JSON file, by dates for the last year, each day contains an array of items

that were sold/purchased on it. For each item we have the item-number, it's price and quantity:



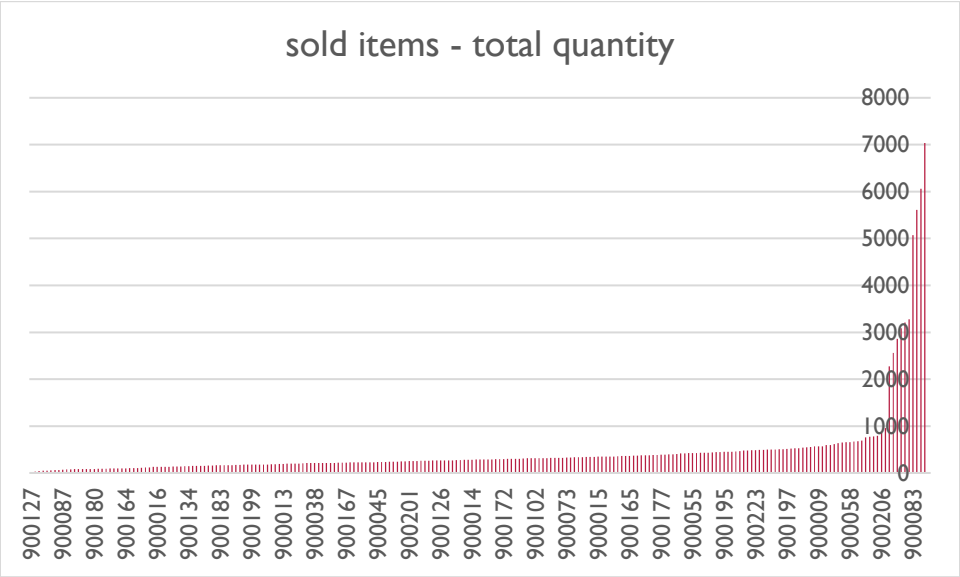
```
{
  "date": "2015-05-04",
  "total": 375,
  "details": [
    {
      "name": "900284",
      "date": "2015-05-04",
      "value": 74,
      "price": 11
    },
    {
      "name": "900272",
      "date": "2015-05-04",
      "value": 37,
      "price": 8
    },
    {
      "name": "900121",
      "date": "2015-05-04",
      "value": 67,
      "price": 9
    },
    {
      "name": "900194",
      "date": "2015-05-04",
      "value": 28,
      "price": 23
    },
```

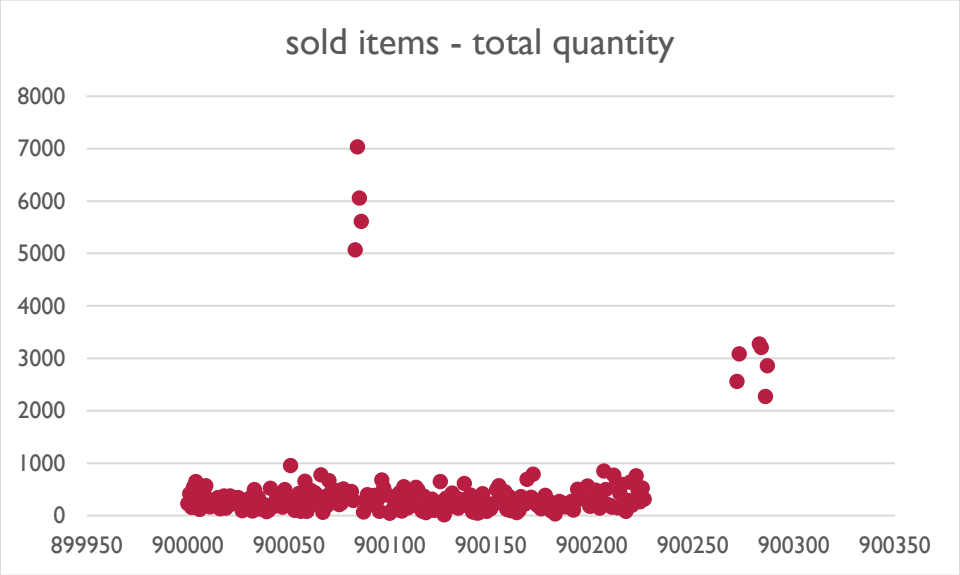
We also used a JSON file that contains all the details about those items:



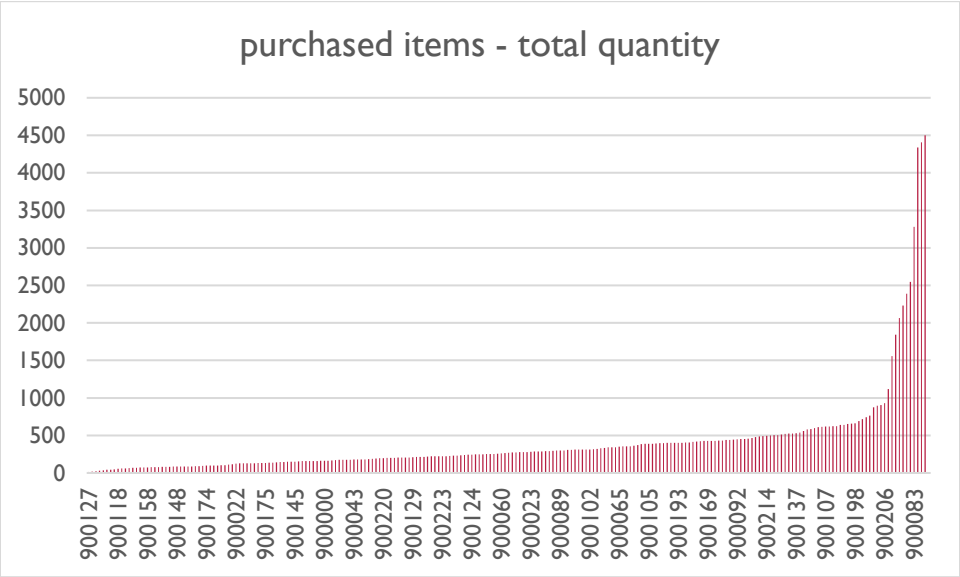
```
[
  {
    "item_number": "900000",
    "item_description": "מ"\ס 3 טבעות פוליו 2 אוגדן-( POF ), צבע אדום עברית",
    "item_unit_price": 11
  },
  {
    "item_number": "900001",
    "item_description": "מ"\ס 3 טבעות פוליו 2 אוגדן-( POF ), צבע שחור עברית",
    "item_unit_price": 11
  },
  {
```

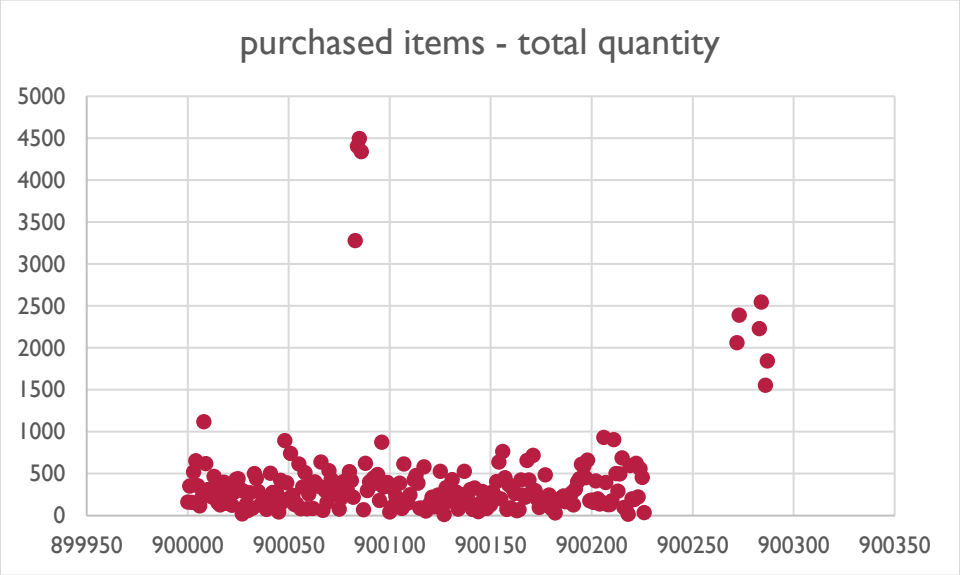
The items statistics over the sales:





The items statistics over the purchases:





*the purchases quantity is lower than the sales quantities because of the frame of the year.

USER'S TASKS:

In this visualization we want to achieve these tasks:

1. To explore patterns of sales for example:

- Specific periods or specific occasions in which there are more sales.

- Discover if a specific item is a best seller product (periodically or all-year). Explore what is the influence of an item in each day by its quantity or by its total cost (quantity*price).

- Comparing sales of two specific products.

- Explore days in which the sales contained high total cost but low quantity of products, and opposite (low cost for a lot of products – high ratio).

2. Explore patterns of purchases for example:

- Explore usual and normal purchases patterns.

- The ability to reveal a fraud or unusual behave at the purchases of the business.

- Discover what is the influence of an item in each day by its quantity or by its total cost (quantity*price).

- Explore days in which the purchases contained high total cost but low quantity of products, and opposite (low cost for a lot of products – high ratio).

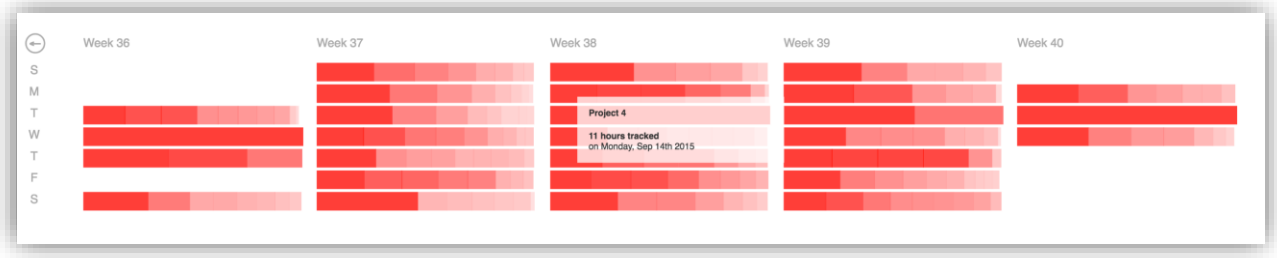
To summarize the tasks, we want to explain the data. The user already knows his data, we want to explain what comes up from the data, to identify outliers and to simplify it especially because of the big amount of data (looking at the rows data wouldn't help at all).

VISUAL MAPPING:

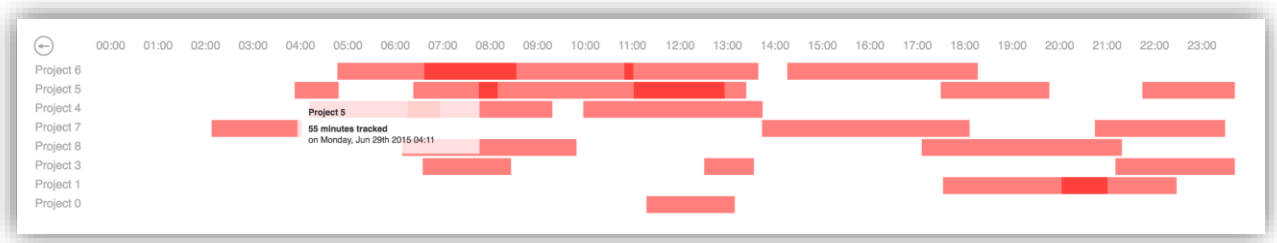
At the beginning of work, we've thought about the three options to visualize time data and we've picked the calendric. We started with searching existing projects and we've found a visualization of a calendar heat-map on github (https://github.com/gleb/angular-calendar-heatmap). We liked this visualization most because of the use of circles instead of rectangles as at the standard calendar heat-map. With the circles it's easier to see the differences of size without the angels and there is an option to use a stroke which is easier to see the ratio with. The example data was Time series data for each day going 1 year back. Each day has a total time tracked (in seconds). Details, if provided, are shown in a tooltip as well as in daily overview. The visualization has different options of overviews:

Year overview

Month overview



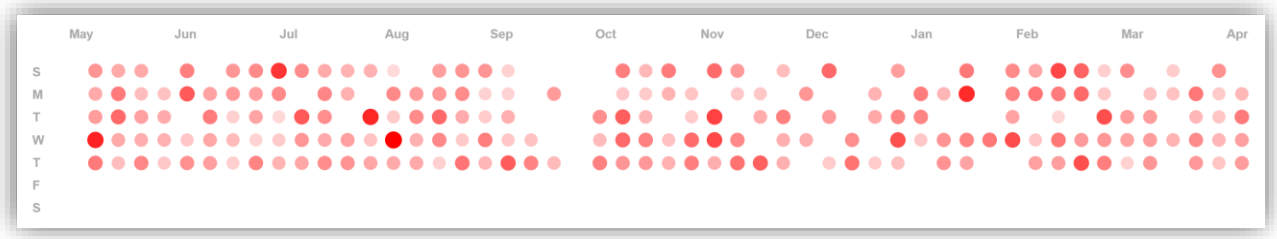
Daily overview



We started to change it and to adapt it to our needs. First we've changed the 'project' to be the 'item number'. Then, we've changed the 'duration of project' to the amount of this item and the 'total of duration' to be the total of all the sales or purchases at that day.

First step – Adding our data (sales) to the visualization:

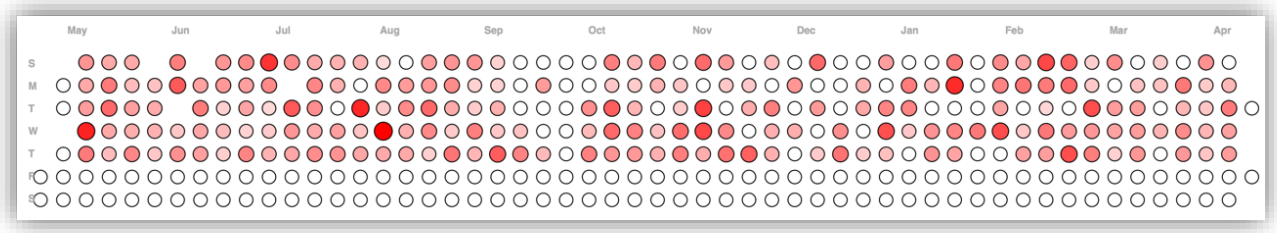
We adapted the visualization to our attributes and we took off the overviews that didn't fit to our needs for example the daily and weekly overviews.



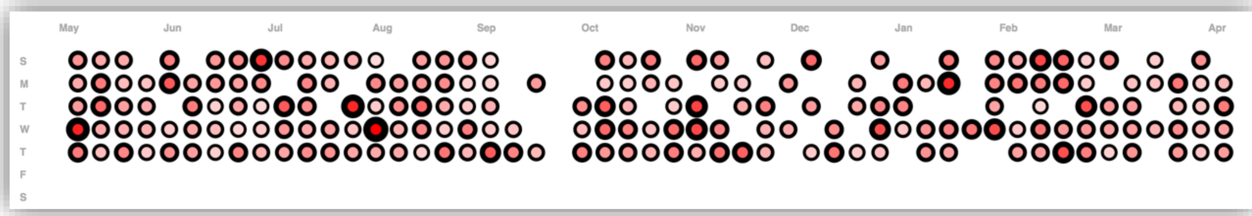
At this point we figured we needed to work on our data and to proceed with the work on the visualization itself. The hex of the color and the size of each circle represents the total quantity that was sold (at this example for sales) that day compared to the rest of the days.

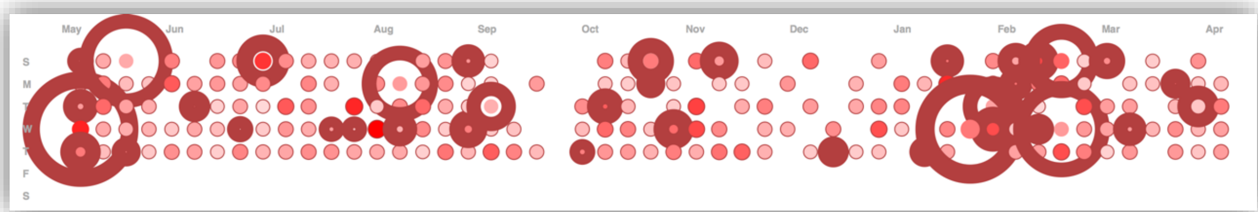
Second step – The ratio of one item to others:

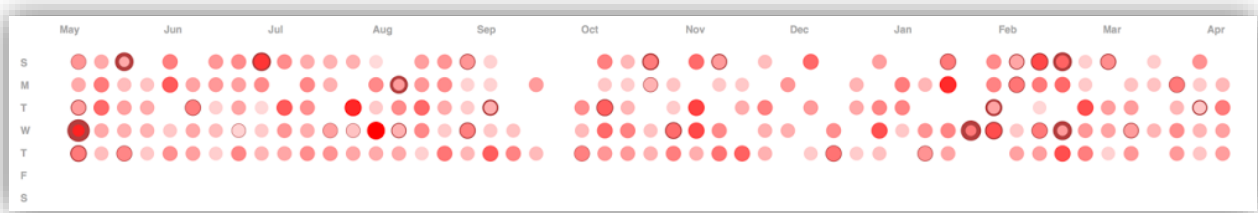
We wanted to show the quantity of a selected item compared with the total sales quantity of each day this item was sold. For that purpose, we've used the Stroke. We added a dropdown that contains the items that were sold, and after selecting each item the calendar heat-map refreshes and shows the stroke of each day by the percent of the selected item compared to the total of the other items. Those are the technical steps: We added the stroke,

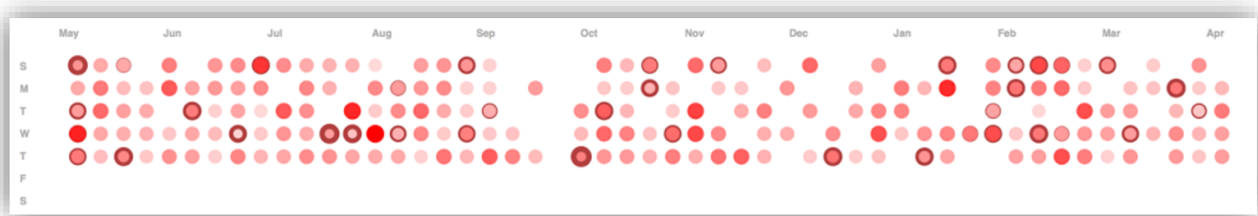


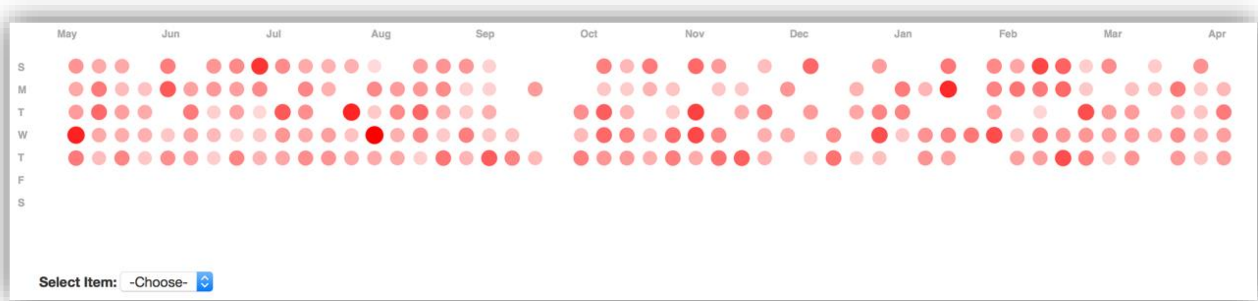
We took off the stroke from the days that the specific item was'nt sold on (and days that didn't contain sales at all), we chaned the weight of the stroke and tried to create it based on the ratio.


May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr
S M T W T F S

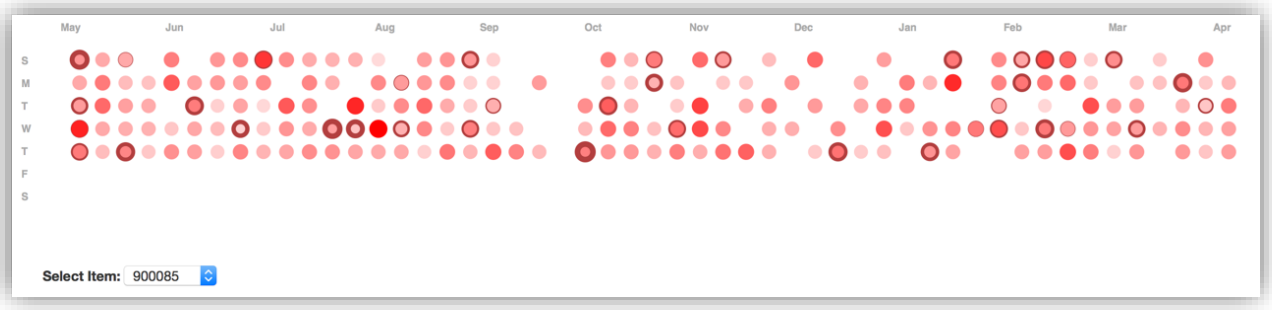

May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr
S M T W T F S


May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr
S M T W T F S


May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr
S M T W T F S

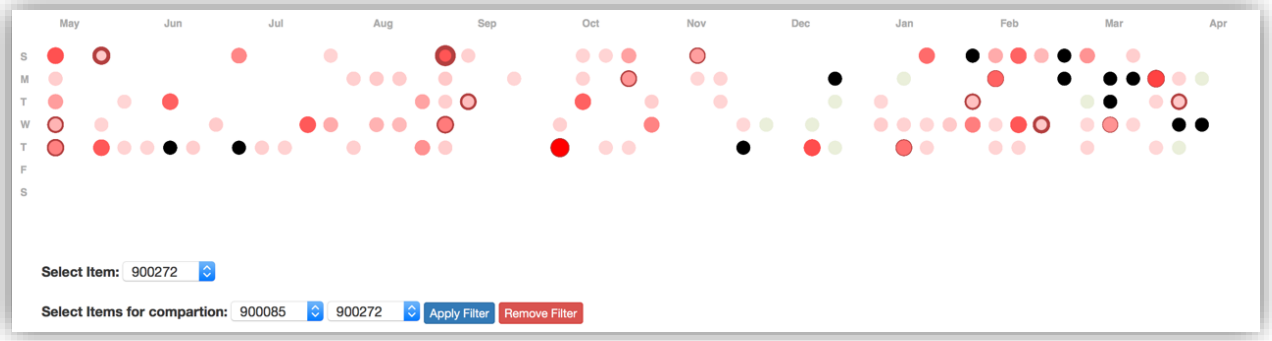

May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr
S M T W T F S
Select Item: -Choose-

After choosing an item:

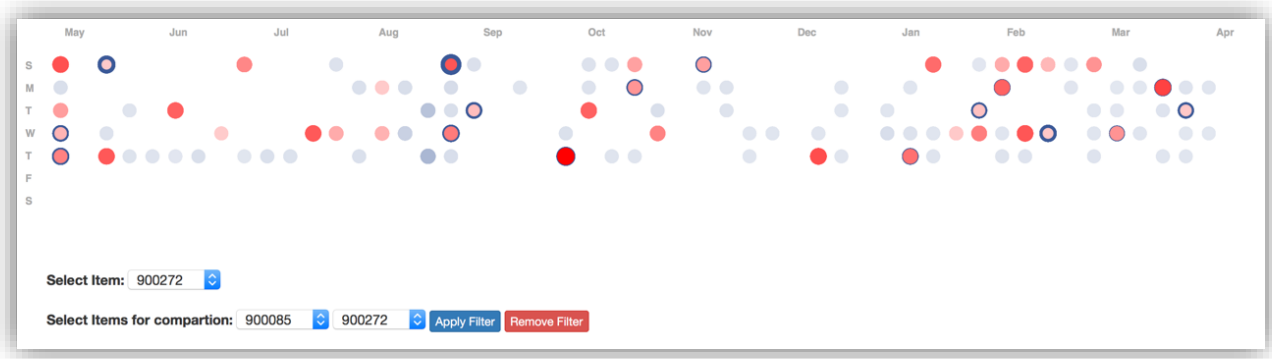


Third step – The comparison between two different items:

At this point we wanted to create the feature of comparison between the sales of two different items. We created two dropdowns with the items and made a filter which shows the sales details for those two items only.

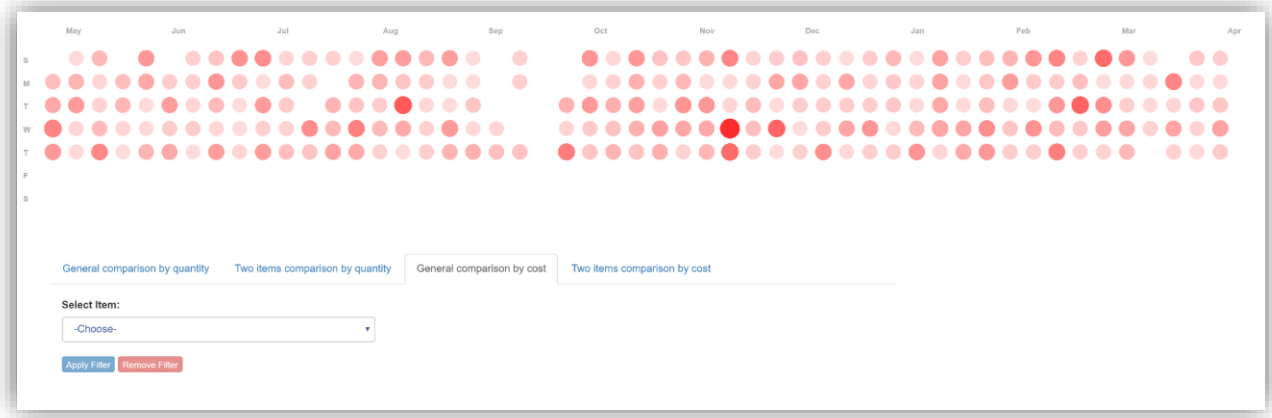


It didn't show it properly and it wasn't a good combination to show the differences so we changed it to be red and blue (and the stroke to be blue anyway). At the days in which only the first item was sold (without the second one), it would appear like at the normal view (red color) only. At the days in which only the second item was sold (without the first one), it would appear in blue color only. At the days that the two items were sold, the stroke's weight would represent the ratio of the second item compared to the first one (divided by the total of both of them). To return to the first view without no filter the user can click on "Remove Filter".

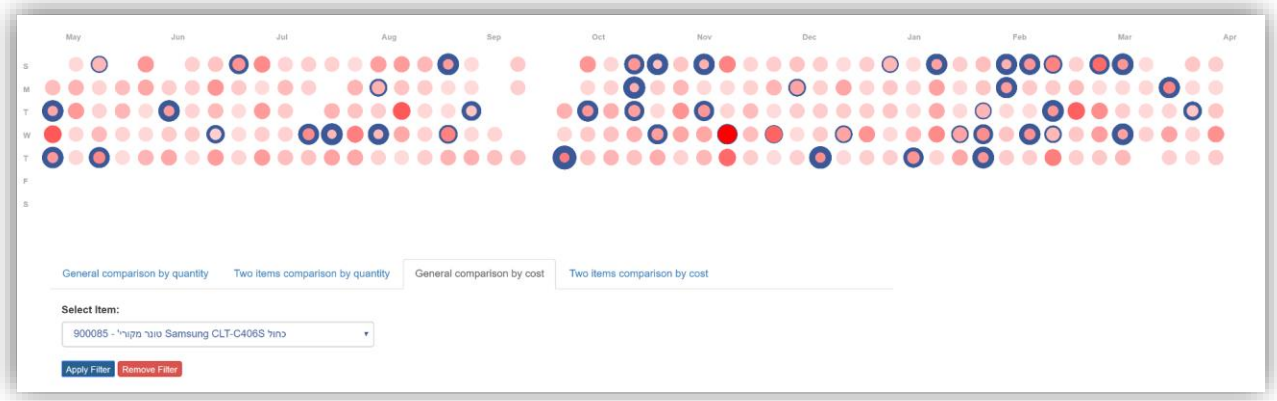


Forth step – The comparison between to different items by cost:

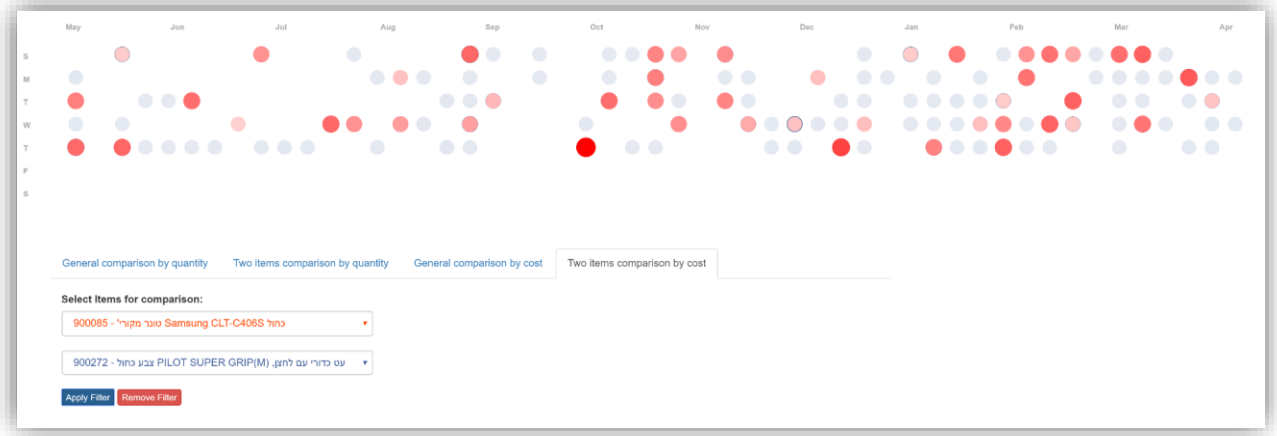
After seeing the differences between items by their quantity, we realized it wouldn't be enough to show the quantity only because an item which was sold with a big number of units could have a low price and would be a little part of the total cost. Therefore, we added an option to compare two items (or just one using the stroke only) by the percent they represent of the total cost on a specific day. Now we have four different filters so we've put them in tabs to make it more comfortable and intuitive.



The comparison by a specific item (by its cost compared to the total cost):



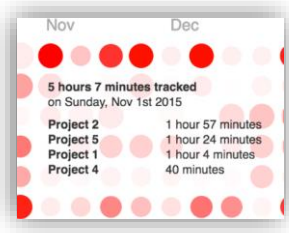
The stroke represents the percent of the total cost for this item. When we want to compare to different items we can use:



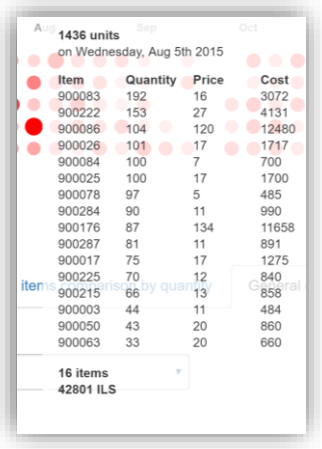
The colors and the stroke are the same idea like the quantity comparison but with the cost.

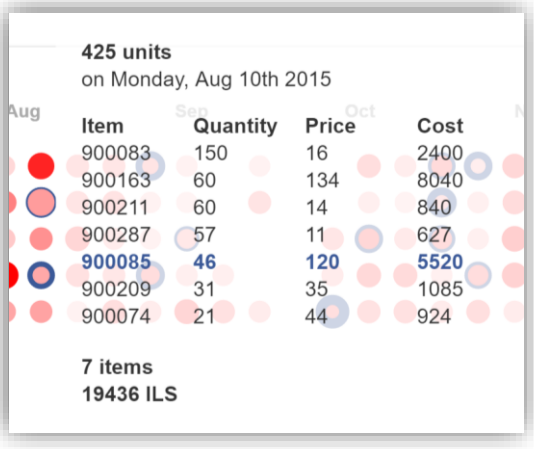
Fifth step – The tooltips:

At the original visualization each circle represented a duration of projects:



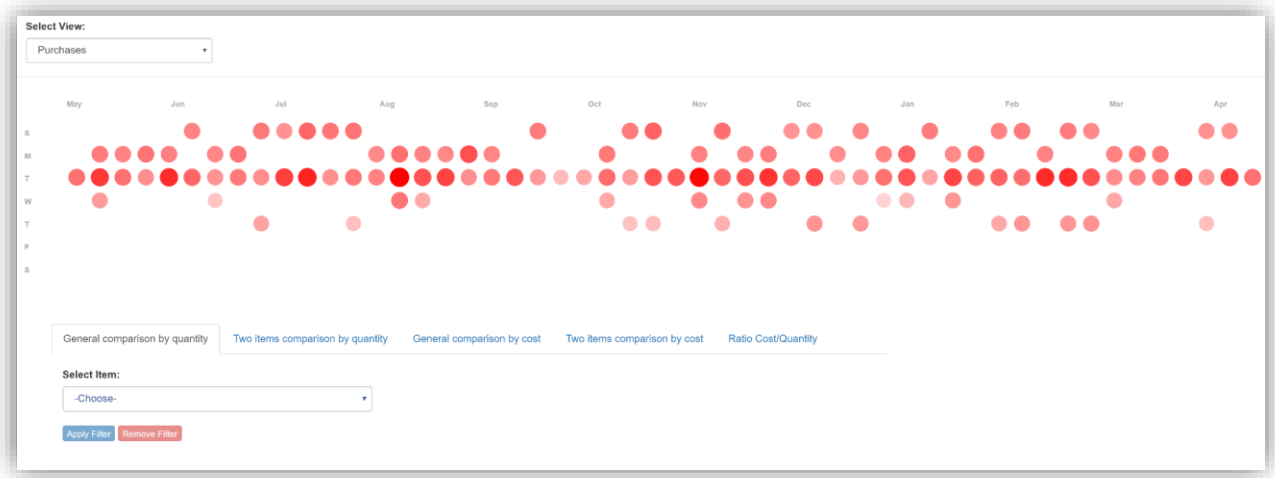
In addition to the first changed we made ('project' to 'item number' and 'duration' to 'quantity'), we added a total quantity of the sold (or purchased) units, the price for each item, the total cost for each item (quantity*price) and the total cost for each day (plus number of different items). The chosen item for the stroke would be bold and blue at the tooltip (like at the second image):





Sixth step – Adding a purchase-orders view:

We added a dropdown with 2 choices – sales or purchases. After clicking on purchases the view changes and contains the purchase-orders data. It's the same idea as at the sales.



Seventh step – Adding a filter for Ratio of cost/quantity:

We added a tab for this filter which shows the ratio of the total cost divided by the total quantity for each day (sales or purchases).

To summarize this section, we've tried to complete each and one of the user's tasks. The calendar heat-map gives the option to see the data in an intuitive and simple form, it's easy to find patterns like that by the size and the hex. By using the filter, the user could find out what is the influence of an item on the total quantity or cost (either the sales or the purchases) etc.

VALUE OF VISUALIZATION

a. V = T + I + E + C

Time weight = 0.3 Insight weight = 0.3 Essence weight = 0.3 Confidence weight = 0.1

Time: we've made the data more clear and simple to read by using the calendar heat-map and filters. It's more visible to a human eye to realize what differences and to explore patterns very quick.

Insight: the use of the color hex and the calendric form contribute insights about different periods over a year. The option to filter (using different color or stroke's size and color) by an item or to make comparison between to different items gives the possibility to get very quick and good insights about each item.

Essence: we've tried to give the user an option to get the overall essence by looking at the circles which represent each day over a year but we also wanted to give the user more important details about each day and for that we've used the tooltip (items sold/purchased on each day and their quantity, price and cost, and also the total cost and total quantity per day. The filters and the option to highlight weekdays or months are giving the user more abilities to get sense of the data.

Confidence: the details we've showed on the tooltips with the exact quantities and costs (plus the totals and calculations) help the user to see that the sense he got and that the algorithms behind the visualization are true and reliable.

b. Review of Design:

In our visualization we've used some attributes to make it more simple and visible so that the user would get his answers as quick as possible.

Pros:

- ✔ The circle for each day on a calendric view simplifies the big amount of data for a human eye and helps the user to get quick conclusions about different periods. Moreover, it adds the option to use a stroke to represent different things.
- ✔ The hex of color and the size were used to make it more clear and to emphasize the differences between the days. Moreover, it helps to highlight patterns like different periods over the year.
- ✔ The option to highlight a specific weekday or a month over a year helps to see those kind of patterns too.

✔While filtering by an item, the stroke's size helps to realize what part it takes of the total cost or quantity at each day. It's easy to note it because of the different color of the stroke. Like that we could realize if it's a best-selling product.

✔While comparing two different items it's easy to see in which of the days each one of them was sold/purchased without the other one, in which days they were sold/purchased together and what was the difference between them in terms of quantity and cost.

✔The tooltips give the user a very essential details like the items that were sold/purchased on each day and their quantity, price and cost, and also the total cost and total quantity per day.

✔The ratio tab gives the user the option to find days in which he sold/purchased a lot of products with a low total cost and days in which he sold/purchased a little number of products with high total cost.

Cons:

✘The fact that not all the circles have the same alignment and not all of them are close to each other, makes it more difficult to compare by the size and by the hex of color.

✘It's a little bit difficult to get the exact differences by comparing the hexes and especially when they don't appear aside each other (we can differ by hex only if it's a big difference).

✘The calendric form makes the differences of sizes very limited because it needs to fit the window and contain the whole year.

✘The same thing about the stroke and it's difficult to get a precise notion about the ratio and the differences between the strokes.

✘There is a limitation for the ratio of the stroke and while it is a big percent (above 90%) of the total it could be impossible to see the hex of the circle inside the stroke (it doesn't happen at the comparison by one item, but for the filter of two items it can happen and to fix it we can just change the order of the dropdowns).

✘The comparison by filter is useful only for one or two items, for more it's not effective.

✘Showing the tooltips only while hovering over a specific circle doesn't allow comparison by items without using the filters, there isn't an option to see more than one tooltip each time.

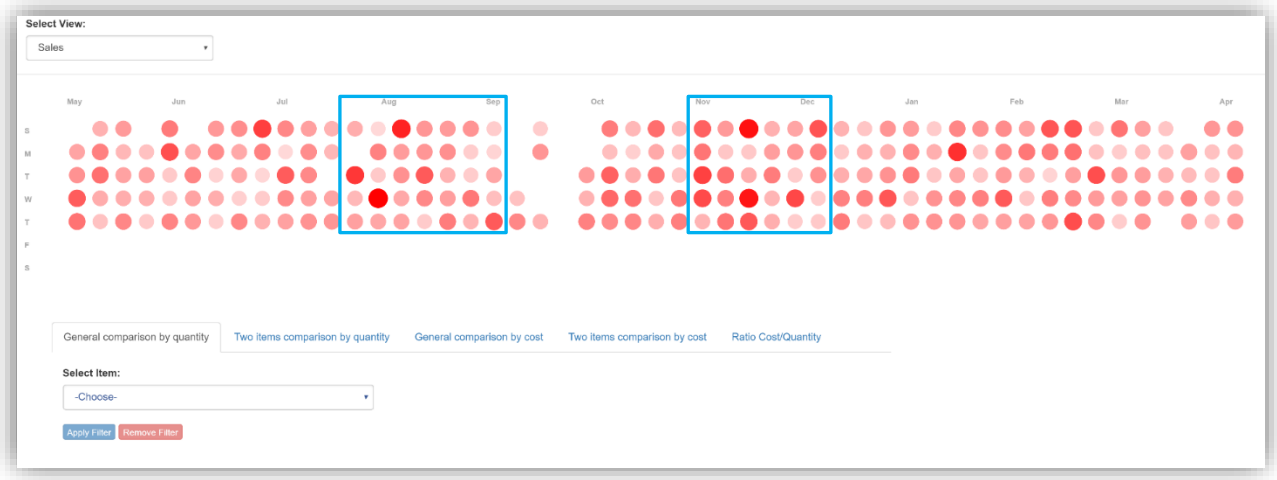
✘With a bigger amount of data in each day it will be impossible to read and to use the tooltips (it will overload the screen but for this business it can't happen).

Examples for user's task answers:

1. To explore patterns of sales for example:

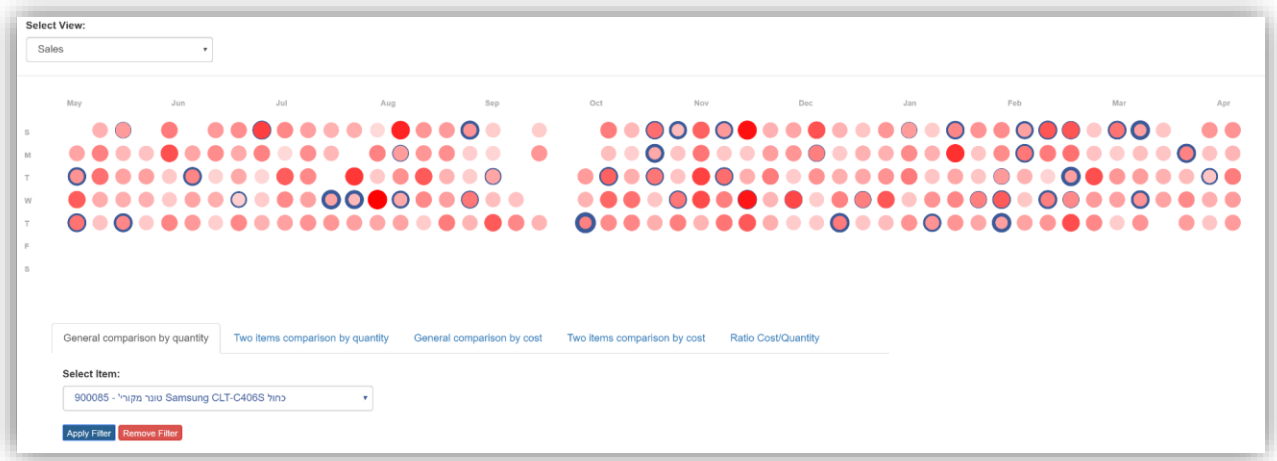
- Specific periods or specific occasions in which there are more sales.

You can see on the default general view for sales (by quantity), that the busiest days are on august (also at the end of July and the beginning of September), which is right before the beginning of school's year. You can also see that at the end of the year, especially at the end of November and at the beginning of December there are a lot of sales.



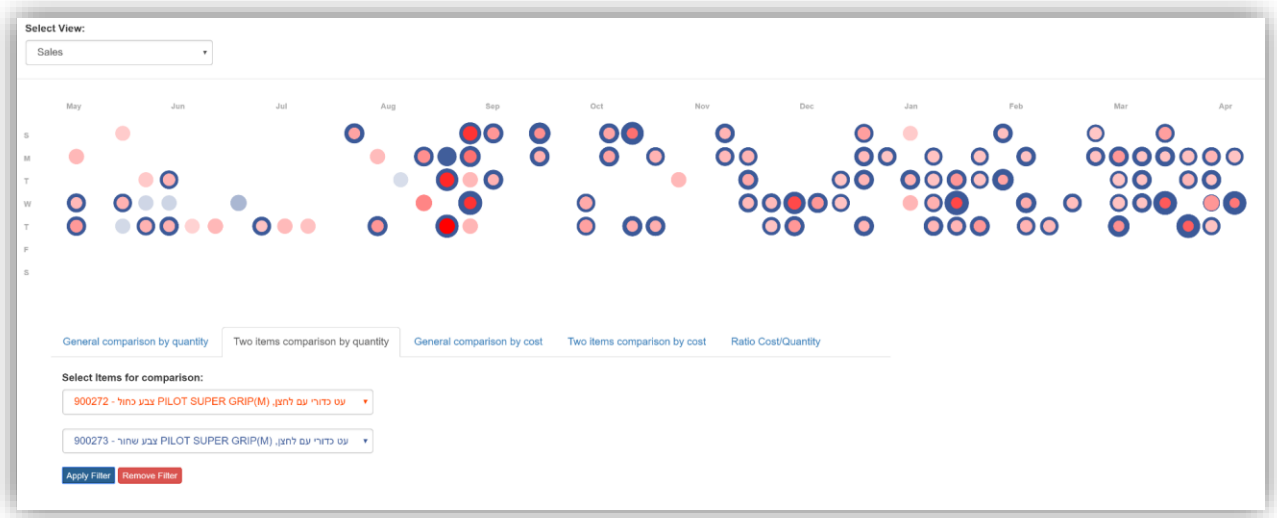
- Discover if a specific item is a best seller product (periodically or all-year). Explore what is the influence of an item in each day by its quantity or by its total cost (quantity*price).

There is an option to see for each item if it's a best-selling product, and to discover by hovering on a month or weekday if it's periodically or not.



- Comparing sales of two specific products.

There is an option to use the comparison filter between two items. You can see the difference by the quantity filter or by the cost filter. Each one of them would appear with the dropdown color and on the days in which they were sold together, the item at the second dropdown would appear as the stroke (it would also appear at the tooltip in blue). In this example we can see that both pilot super grip – black and blue are usually being sold together and almost at the same quantities.

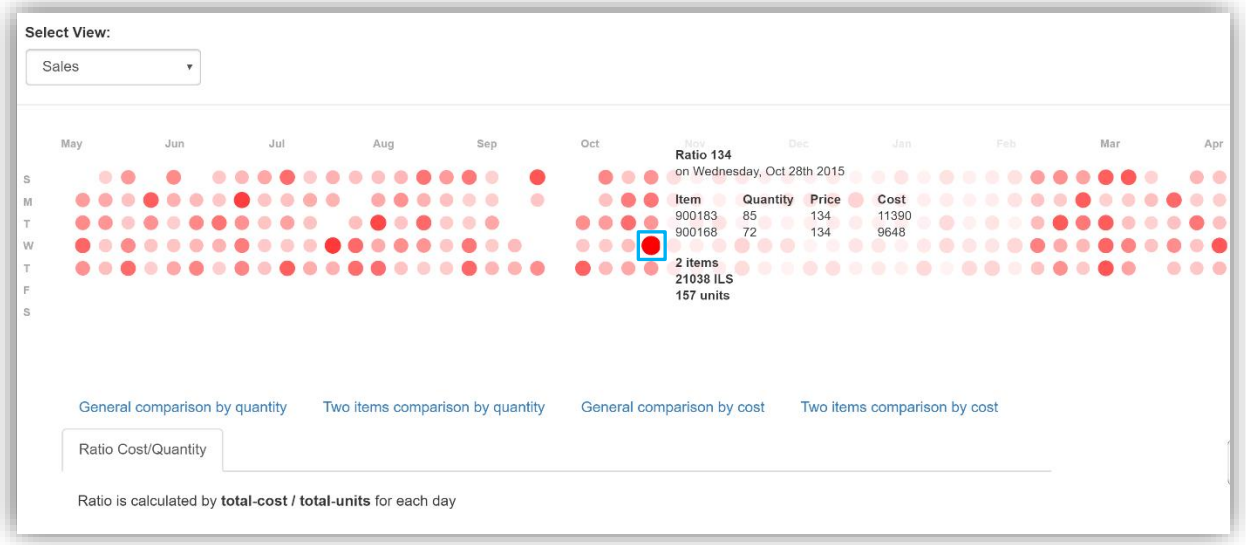


- Explore days in which the sales contained high total cost but low quantity of products, and opposite (low cost for a lot of products – high ratio).

On a day in which it appears that it's not a busy day by the quantity general view (sales), we can look at the ratio tab and discover that there was a little quantity sold but with a big amount of money, for example:

Sales view on the quantity tab:

Sales view on the ratio tab:



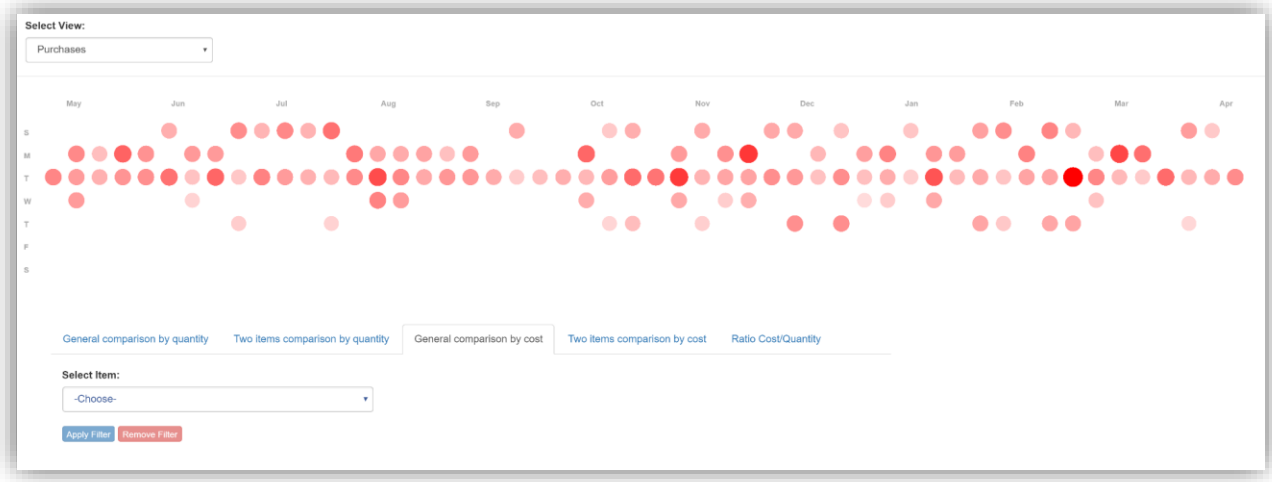
2. Explore patterns of purchases for example:

- Explore usual and normal purchases patterns.

By hovering on Tuesday we can see that most of the purchases of this business are made on the middle of the week. Almost in each month there is a big purchase (at most two), except the holidays (September – October).

- The ability to reveal a fraud or unusual behave at the purchases of the business.

After we know the usual patterns we can discover unusual behavior and like that we can detect fraud. We can also look at a specific item and to see if there are unusual purchases of it. The highest total cost at this year is 53,724 ILS at Mars (which is ok according to the owner explains). We can see that there aren't outliers and the owner could discover a fraud that would be more than this amount of money.



- Discover what is the influence of an item in each day by its quantity or by its total cost (quantity*price).

(The same as at the sales example)

- Explore days in which the purchases contained high total cost but low quantity of products, and opposite (low cost for a lot of products – high ratio).

(The same as at the sales example)