Final Report - VAST Challenge 2017

Bianca Barnucz *
Environment Reseacher of Mistford

Stephanie Wegscheidl †
Environment Reseacher of Mistford

ZUSAMMENFASSUNG

In the last weeks and months we have been working on the VAST Challenge of this current year and developed through a design study various Tableau dashboards which present the given data and should assist users to understand and comprehend the recent changes within the preserve. We have chosen this kind of topic for our group project because we do believe that the theme of this VAST Challenge is really interesting and exciting to visualize. The VAST Challenge consists of three so called Mini Challengesänd one final Grand Challenge. Especially these different Challenges appealed to us and we hope that we could develop suitable visualizations for the given problems. In the following paragraphs and pages you are going to read about the tasks we had to accomplish, the given data, our approaches, the implementation process, difficulties we had to deal with, results of our work and as well as about the strengths and weaknesses of our visualization systems.

Index Terms: VAST Challenge 2017—Visualization—Visualization techniques—Tableau;
Mistford—Birds—Mitch Vogel—Environment

1 MOTIVATION

1.1 Information about the Problem

This year's VAST Challenge was developed to help the post-doc student Mitch Vogel who discovered signs that the number of nesting pairs of the Rose-Crested Blue Pipit, a rare species of bids, is decreasing in the city of Mistford. Furthermore, these VAST Challenges deal with the environmental changes of Mistford's nature park which might be the reason why the population of the birds is decreasing. Mitch Vogel has received various datasets which he wants to analyze in order to pursue further steps to save the Blue-Pipit's habitat. To help Mitch to figure out what, who, where, when and how the situation within the preserve has changed various Challenges have been designed. Therefore we have created four dashboards which bring different aspects of the data into focus.

1.2 Challenge 1

1.2.1 Task

Mitch has notices some untypical behavior of the vehicles visiting the park. He believes that the noise and the pollution of the traffic could be one possible aspects which might disturb the birds in their nesting process. To observe the behavior of these vehicles which are visiting the park Mitch would need a visualization of all the collected data from the sensors which are monitoring the entrances.

1.2.2 Data

For the first Mini-Challenge we received a word document with the data description and two images showing the roads of the reserve. In addition, we also got a CSV file which describes the results of the sensors which measured the behavior of the vehicles.

*e-mail: a01501462@unet.univie.ac.at
†e-mail: a01502707@unet.univie.ac.at

1.3 Challenge 2

1.3.1 Task

Nearby the preserve there are four factories which are harming the environment by pouring noxious gases out of their smokestack. These gases might have quite a big impact on the decreasing population of the birds and their health. However, Mitch knows that these companies are really trying to be more environmental friendlier and are very ambitious to follow the regulations concerning the toxic gases. Mistford has been monitoring the air around the factories over three months and Mitch was able to get some insight of the data. Even though Mitch is an expert in Excel he believes that there are better ways to present the data so that it will be easier to analyze. By creating a visual analytics of the current situation concerning the factories Mitch would get a better overview of the environmental assessments surrounded by the preserve.

1.3.2 Data

The data for Challenge 2 includes two Excel files with meteorological and sensor data. The sensor data describes the measurements of the different gases while the meteorological data presents the wind directions and the wind speed. Furthermore, this dataset includes two word documents with detailed descriptions about the factories and the sensors of the Preserves area. Last but not least it also includes a map of the Preserve with labels on it.

1.4 Challenge 3

1.4.1 Task

As Mitch is working independently and due to the size of the wildlife preserve he is not able to cover all of it and to keep an eye on the environmental impacts and changes. To better understand the preserves health and to alert him when the health of the preserve is changing he would need a visual analytic approach. Moreover, he has also talked to a professor who supplied him with images of the National Space Service. Therefore, Mitch needs someone who analyses these images and draws conclusions out of them.

1.4.2 Data

The data for the third Challenge includes an image of the Boonsong Lake and a description of the Lekagul preserve. Moreover, it also provides a total of twelve images of the nature park and twelve CSV files with additional information about the preserves area which were measured by sensors. Furthermore, it also provides a description of how to use the multi-spectral images and its data.

1.5 Grand Challenge

1.5.1 Task

All three visualizations which were done earlier are very important to Mitch to analyze the behavior of the vehicles visiting the park, the chemicals from the factories as well as the health of the preserve. Nonetheless, it is not only important to analyze what is happening in the preserve, but also to know what, where, why, when, who and how changes have been taken place. Therefore, he would need another analyzation which combines all of the previous information and visualizations and shows the connection as well as the relations between the individual views to understand the Wildlife even better. Hopefully this final visualization can help him to pursue further steps to save the Rose-Crested Blue Pipit.

1.5.2 Data

The dataset for the Grand Challenge consists of various Mistford Industrial Park newsletters from different months between the years 2013 2016. These newsletters might help us with the final report and to draw connections between the previously used datasets.

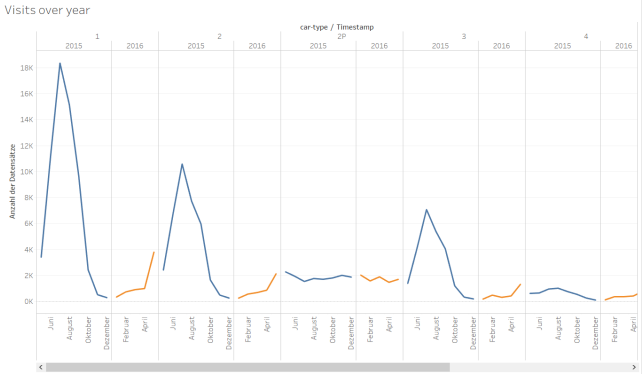
1.6 User

For our project we have one main user, Mitch Vogel, a post-doc student form Mistford. He wants to save a rare species of birds called Rose-Crested Blue Pipit. He needs all of the previously described visualizations to get a better understanding of the Wildlife at the preserve and an insight of the reasons how different aspects are influencing the bird s population. Furthermore, he needs all the conditioned data to pursue further steps if needed to save the birds. Even though Mitch is our primary user, other users such as professors or rangers might profit from our visualizations.

2 Related Work

It is always interesting to compare its own work with different approaches and techniques others have used. Since a couple of people participated in the VAST Challenge 2017 it was not that difficult to find other reports and visualizations of the given tasks. We took a few examples and compared them to our approaches and solutions.
For example we have found a document in which another team had similar ideas that we had for interpreting the first Challenge - to show the general gate use. The graph bellow states that there is a seasonal increase of the gate use during the summer months, which means that more visitors are entering the park in summer. We think that the other team has chosen a quite complex and maybe even distracting way of showing a pretty simple subject. [1]

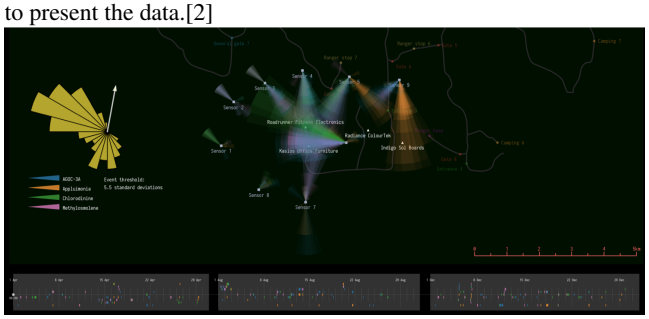
We believe that we have found an easier way to represent the data. However, we used a line chart to show when the most visitors entering the park can be listed during the year.



In comparison to the other approach, we even gave additional information about the car-type as it can be seen in the image above.

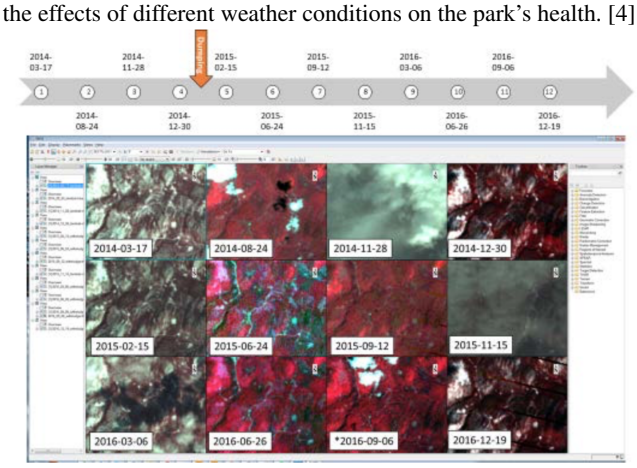
Even though, the visualizations from the other team might look more professional and high sophisticated, we are convinced that we did provide a better insight into the dataset due to the integration of filters which haven't been used in these other visualizations.

On the other hand a group from the University of London created a very detailed and interesting visualization for Challenge 2. Through these visualizations they were able to draw more exact conclusions about the pollutants which are affecting the nesting habit of the birds living within the nature park. In comparison to our approach this team has used another visualization tool then Tableau to present the data.[2]



What we really adored about their approach was the way they realized the presentation of the sensor's data and the wind directions which really helps the user to understand which factory is responsible for the release of which kind of chemicals. Furthermore, just like we did, they emphasized their visualizations through interactions. Instead of filters they applied another form of interaction between the various views as it can be seen in the video they have made. [3]

Moreover, another team which dealt with the effects of the environmental changes, had very alike ideas about the presentation of Challenge 3, then we did. They also used heat-maps to present the effects of different weather conditions on the park's health. [4]



Even though our approaches might be similar, our implementation, especially the color encoding, differs. This team used black, brown, white and red colors to demonstrate the changes that have been taken place, while we have used the color green to highlight a healthy and densely populated vegetation and gray for the rare vegetation spots. As their choice of color is not stated in the report we do not exactly understand why they have chosen this color combination, as it does not look very appealing. [4]

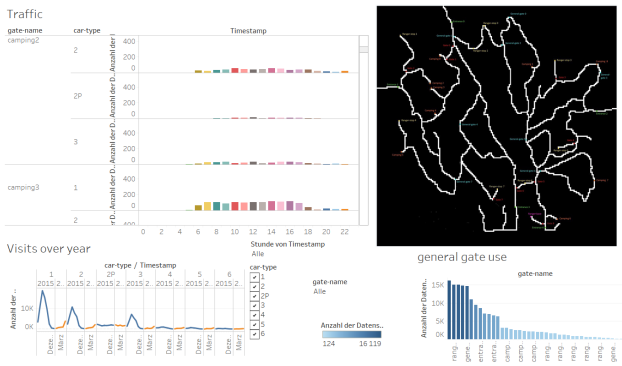
Summarizing it can be said, that even if the approaches and graphs from the other visualizations might have looked very fancy and detailed at first sight, they were not always better or more accurate then ours, simply different. Nonetheless it was exiting to see how many different approaches there were and in what context

the data could be displayed.

3 Approach

In the following chapter we will be justifying our approaches and solutions we have come up with, in order to develop our final dashboards. Furthermore, we have added screenshots which should emphasize our reasons for creating the different graphs the way we did.

3.1 Challenge 1



3.1.1 Description of the Dashboard

The most important view of this dashboard is located in the upper left corner. The user is able to analyze which car-type uses which gate at what time. So it is easy to see which gates are used frequently and at what time vehicles are passing these gates.
For the first prototype we thought it would not be necessary to provide a picture of the preserve‘s map. After having a closer look at the real data we thought it might be easier for the users to combine our visualizations with a map to understand where the cars have been actually driving. In order to understand these routes, we decided to supply a map of the preserve for the users so it is comprehensible which areas are visited the most. Additionally, this map provides an overview where the gates are located. Through the combination of this map and the first graph users should be able to locate where and which car-types has passed the park.
The next view in the lower left corner shows how many visitors have been to the park each month. For this view we compared the numbers of each month form the years 2015 and 2016 to see how the numbers of visitors have changed over the past year. Last but not least the view in the lower right corner states how frequently the different gates are used
The interactions are intuitively and simple. We added filters, so if one chart is clicked the suiting charts filter the data according to the clicked one as well. Furthermore, we added tooltips and sorted the “general gate usegraph from the most used to the less used gate. We also inserted a control-panel“, where users can filter all connected views. Moreover, we provided a filter according to the hour(Timestamp) and gate-names with a dropdown field and car-type(1-6 and 2P) with a checkbox list. These filters enable to get a better overview of the dataset.

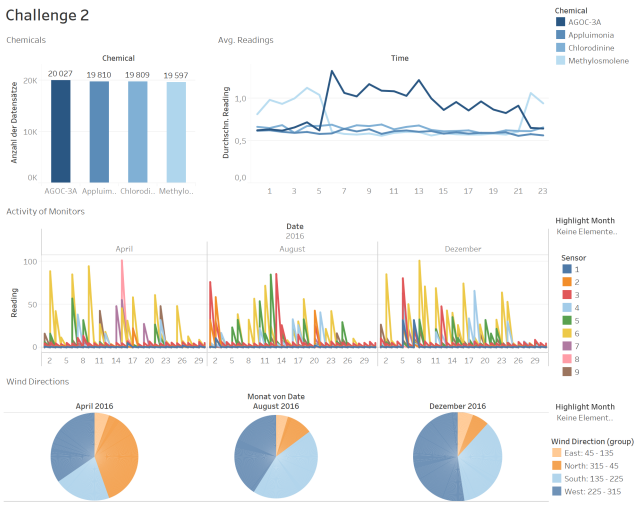
3.1.2 Reasons for our Design Choices

For the dashboard of Challenge 1 we decided to present the most important information from the dataset.
The first view is a bar chart, from which information about the usage of the different car-types and gates over time can be extracted. For a better overview we decided to use different colors for each hour. Furthermore, we implemented a slide bar in this view to get a closer and better insight into the data.
In our opinion a bar chart was the best way to visualize the large amount of data. However, we also thought of creating a line graph for this view but this would have only showed the daily occupancy rate of the park and this was not our aim for this view. Besides, we provided several filters to give the user the possibility to have a more detailed look at the data as it is easier to interpret the data, when the focus is set on one special car-type, gate-name or hour.
Furthermore, we provided a map of the area which shows the location of the roads and the gates within the park. With the help of the data from the first view it is now easy to recognize the areas with heavy traffic.
We also decided to give the user an overview about the general gate use of the park. Therefore, we have made a bar chart which is sorted from the most used gate to the least used gate and we added a tooltip to provide a better overview. The filters we placed on the dashboard can interact with this view as well. For a clear structure we sorted the sum of the data from the highest to the shortest bars.
Last but not least we created a line chart to show the visits over the years. Furthermore, we separated the years in colors to make it easier to interpret and compare the data.

3.1.3 Results

Examining the different views from Challenge 1 we have found out that the most visitors come during the summer months to enter the park. This is probably caused by the holiday season and the good weather. Since this area is a good camping destination this might be one of the main causes for the increase of visitors in the summer. At this time the traffic and air pollution increases due to the visitors which might effect the birds during their breading and hatching process.
After analyzing the data we also figured out that the most traffic is in the areas around general-gate1, ranger-stop 2, ranger-stop 0, general-gate 2 and general-gate 7. While general-gate 7 measured the most passes. Noise and air pollution has to be greater in these areas than in the other ones of the park.
During our research we found out that bigger vehicles, with three axes and more, were recorded at the general-gates 1,2,4 and 7 and at the ranger-stop 0 and 2. This means that these gates were often used by bigger vehicles and have definitely affect the birds. Last but not least we also ascertained that the birds do not have a real resting period where they get enough sleep as there are certain measurements of traffic during the night hours. However, between 10 pm and 5 am there is less traffic than during the day time.

3.2 Challenge 2



3.2.1 Description of the Dashboard

This dashboard represents Challenge 2 and consists of the following four single views: Chemicals, Average Readings, Activity of Monitors and the Wind Directions.
The first chart gives the user an overview of the different chemicals which were detected by the sensors. This view includes a tooltip which contains information about the chemical‘s name, the number of records and the sum of readings. Furthermore, the Monitors chart includes the number of records which were measured by the different monitors. For this chart we decided not to show the exact amount of the values through labels because we belief that it would have made the whole view too crowded and unreadable for the users. By hovering over the separate charts a tooltip including information about the sensor‘s number, the detected chemical and the exact amount of the measures will appear. It is possible to have a closer look at only one chemical. To pursue this action, the user has to select one of the chemicals from the legend on the right side. If one chemical is chosen the charts of that certain chemical will be highlighted in the Chemicals, Monitors and Activity of the Monitors diagrams.

The Activity of Monitors presents the three months in which the sensors measured the different chemicals in the air. If the lines do not include any spaces it means that the sensors were running faultless and all measures they have made were recorded. On the first sight this chart might look a little bit crowded due to the different colors which represent the sensors. However, we decided to leave this graph the way it is now because it is important to demonstrate the differences between the sensors. Moreover, if the user would like to check if a certain sensor was working flawless, he/she needs to filter for one sensor from the legend to focus on a specific sensor. If the user wants to concentrate on the recording period of the sensors, then this can be done by selecting a certain month through the highlight menu on the right side. Moreover, as previously shortly described to examine exactly the performance of the different sensors each sensor can be selected separately in order to scan if the sensors were working through the entire time.

The Wind Direction view shows from which direction the winds came that spread the chemicals into the air. Due to the different datasets from which we extracted the meteorological data which we used to present the wind directions, it does not have any common data compared to the sensor dataset. Therefore, we could not include any interactions between the Wind Direction graph and the others. Nonetheless, if the user is interested in only one specific month then he/she can highlight only the one pie chart through the menu on the right side. Furthermore, by choosing one of the wind direction groups from the legend it is possible to highlight the clicked category in all of the pie charts.

3.2.2 Reasons for our Design Choices

For Challenge 2, where we had to examine the performance of the sensors which were measuring the amount of various chemicals, we decided to present the collected data by using different types of diagrams. The simplest way to demonstrate and to process quantitative data such as the amount of readings of the gases is to visualize them through bar charts. For each bar we used the same color but with different hues - the darker the color the more this chemical is substituted in the air. As another requested task was to show that the sensors were working faultlessly during the measuring period we used a line chart which allows the user to examine the exact dates and to see if there has been a erroneous sensor. To express each sensor we worked with different colors. We could have done the same as in the previous chart, to vary the hues, but we think that due to the very similar curve progressions this might have been too confusing and the data would not have the same effect then it has through the different colors. The last chart type we used for this dashboard are pie charts. Even though the radius of a pie chart and the compass angles are not the same, we thought that it would be suitable to present the wind directions through pie charts, which we divided into four categories which stand for the cardinal directions: North, East, South and West. We tried to edit our pie charts according to the feedback we received, but unfortunately, we were not able to divide the sectors into four similar pieces. However, we do think that this charts label the wind directions in an understandable way. Nonetheless, we did redefine the cardinal directions according to the angles of the compass. Due to this transformation it can be seen, that in April most of the winds came from the North, in August from the South, while in December winds from the West were dominating.

3.2.3 Results

For the second Challenge it was important to detect the sensor’s performance if they were working flawless over the measuring period. Moreover, we had to examine which chemicals were released into the air and if something untypical or conspicuous could be detected.
Our original plan was to be able to make possible predictions about which company is responsible for the release of which chemicals in the end. To be honest we thought that it would be much easier to prophesy which factory generates which chemicals by combining the data from the Monitors“view and the sensor’s locations which were marked on an map.

Furthermore, as we were looking to find some answers to the requested questions we realized that we did not have any chart in our dashboard which we could use to make assumptions about abnormal activities in the context of the chemicals. However, we replaced the Monitors“graphic through a new one which states the average readings of the chemicals on the y-axis and the average hours per day on the x-axis. Through this simple, but still very efficient line chart it is possible to distinguish that even though there have been strict regulations introduced, concerning the release of different pollutants through the Mistford Pact of 2010, the amount of AGOC-3A and Methylosmolene have reached quite critical values. At the same time this newly added chart displays very clearly that chemicals, especially AGOC-3A and Methylosmolene have been released into the environment during very unusual times,

between 11pm and 4am. Although, AGOC-3A is described as a less harmful chemical in comparison to the other substances, it is still effecting the environment and the bird's population. Besides AGOC-3A, Methylosmolene is the more concerning pollutant. As Methylosmolene can cause toxic side effects in vertebrates, and has been strictly regulated in the manufacturing sector. Nonetheless, it needs to bee chemically neutralized before disposal. We belief that inaccurate disposal of this chemical could be the reason why the Lake of Boonsong has changed drastically, as described at Challenge 3.

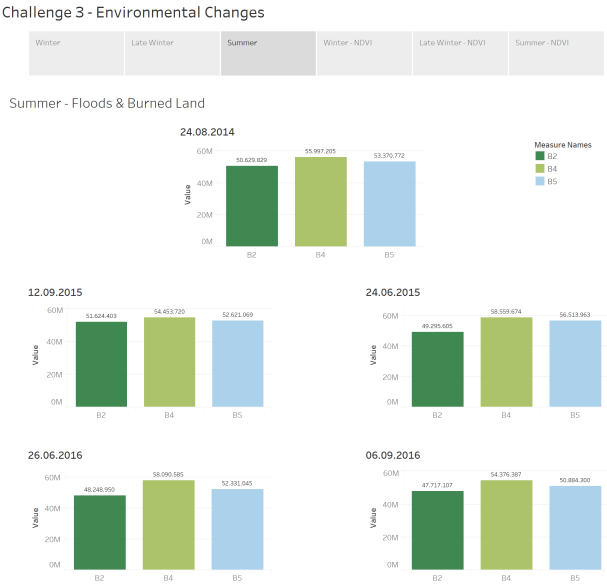
Another interesting feature which can be seen through our visualization are the wind directions from where the winds came. During April and December winds coming from the west were dominating while in August the wind came most likely from the south.

3.3 Challenge 3

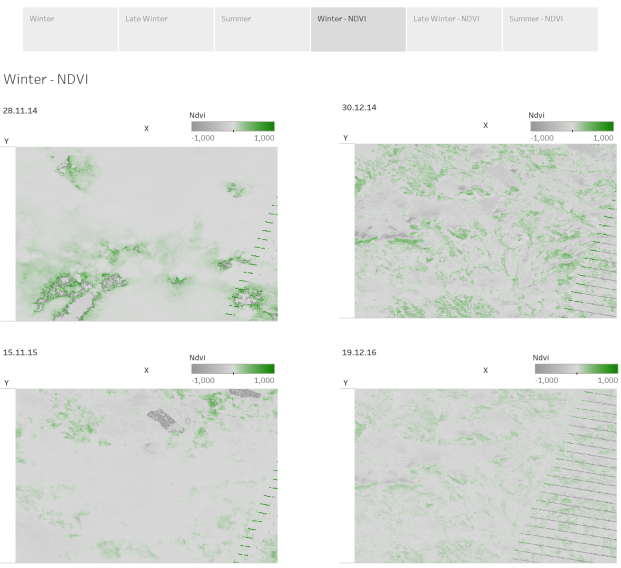
Winter

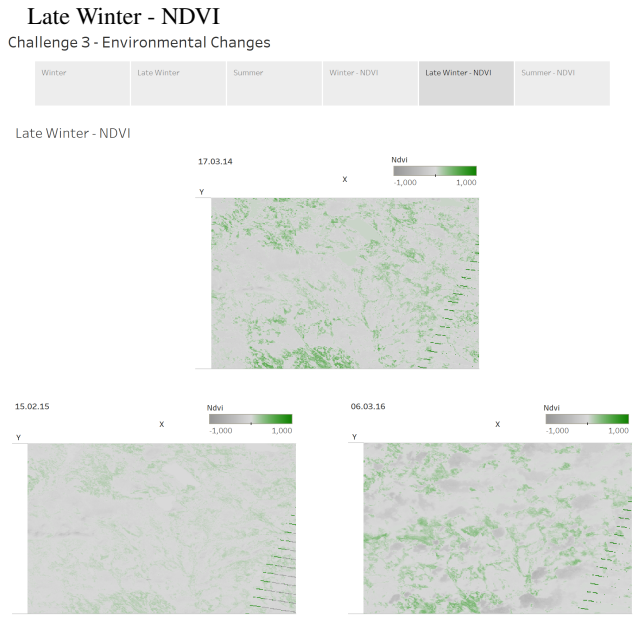
Late Winter

Summer



Winter - NDVI





3.3.1 Description of the Dashboard

For the third Challenge we arranged the different views in form of a story so that the users are able to switch between the different dashboards through the storyline on top of the page. For Challenge 3 we received 12 different pictures and their corresponding CSV files. After having a closer look at the different images we realized that it would make sense to group them by the years seasons: winter, late winter and summer.
However, the first three dashboards (Winter, Late Winter, Summer) include data about the different multi-spectral bands and the color encoding of the images. By combining the three B1, B5, B6 values it can be detected if there has been a lot of snow, ice or clouds at the time when the picture was taken. Furthermore, when the B2, B3 and B4 values are mixed, assumptions about the health of the preserves plants can be made. When joining the B2, B4 and B5 values it can be seen that there have been floods or burned land within the preserve. Through these bar charts users should be able to draw conclusions about the preserves health and condition.
Additionally, the following three dashboards (Winter NDVI, Late Winter NDVI, Summer - NDVI) include information about the Normalized Difference Vegetation Index which can be estimated through the following equation: NDVI = (B4 B3) / (B4 + B3). The higher the NDV-Index, the denser the vegetation is. If the index is rather low it is an indicator for a non-vegetation region. We decided to present these indices for each image through heat-maps which demonstrate the vegetation of certain areas.

Due to the fact that this story view should give Mitch only an overview of the preserves health and environmental changes we decided to spare out unnecessary interactions between the different views. However, we included the option to highlight only one of the B-values by clicking on one of the bars or choosing one from the legend for the first three dashboards. Furthermore, we believe it is important to show within these charts the exact numbers of the values. Therefore, we decided to label each bar with the amount of the certain value.
By hovering over the heat-maps a tooltip with the x and y coordinates and the NDVI information will appear. Even though we did not focus on the interaction between the different views and dashboards for Challenge 3, we put a lot of effort into the color encoding of this story because we are convinced that especially the different colors are highly important in these different charts to demonstrate the environmental changes that have been taken place over the last years.

3.3.2 Reasons for our Design Choices

For Challenge 3 we received the largest dataset compared to the other Challenges. From the beginning it was clear, that we had to create more then just one dashboard to cover all this data. After having modeled all of the dashboards we needed, we had to come up with an idea to combine them into one folder. After having compared various options we decided to present Challenge 3 with its 6 dashboards through a story format. By clicking on one of the boxes on the top of the layout the user is able to switch between the different views.
The first three dashboards contain quantitative data and are visualized through bar charts. The content of these charts should present the amount of various B-values which stand for different multi-spectral bands. As these multi-spectral bands contain values of wavelengths we chose the same colors for the bar charts as they present in the color spectrum. However, there were some B-values which do not present a specific color but infrared waves. In these cases we looked up the values in the table which labeled the multi-spectral bands, wavelength and additional information for the encoding and could extract through some useful information which helped us to chose colors that would present the data in the most suitable way.
The last three dashboards show different areas of the preserve. The main aim of these views was to illustrate the environmental changes over the years. In order to demonstrate clear transformations between these pictures we decided to use heat-maps in which we applied the colors green to show densely covered land and gray to display rarely grown areas.

3.3.3 Results

For Challenge 3 we examined in total 12 pictures which were taken previously and their corresponding data which we received in the form of CSV files. Based on these information we created visualizations which can be used to understand the environmental changes within the nature park. Through the dashboards from Challenge 3 it is possible to draw conclusions about the changes of the nesting habit of the Blue-Pipit due to the weather conditions,

plant health and environmental changes.

Late Winter
The late winter charts includes data about the B2, B3 and B4 band values through which users are able to understand the plants health better. B2 shows different types of plants and is responsible for visible brightness in the pictures which indicates that it might have been quite sunny at the time it was taken. Furthermore, B3 contains red color values which stand for the vegetation and certain mineral deposits. Last but not least B4 is a near infrared value through which partially absorbed water can be examined. Furthermore, B4 is sensitive to the vegetation's structure and chlorophyll. However, having a closer look at the measurements from the years of 2014, 2015 and 2016 it can be said that the plant's health did not change enormously over these three years. Nonetheless, all of the three values increased slightly which means that there are various plant types located within the nature park and certain mineral deposits can be found as well. The increasing chlorophyll values might indicate that during the late winter months there was enough sun and light for the plants to operate photosynthesis and to grow.

Winter
The winter diagrams present the B1, B5 and B6 multi-spectral values through which the weather can be detected in a more specific way. B1 penetrates water, shows thin clouds and is responsible for general visible brightness which could be an indicator for snow. Furthermore, B5 could point out the amounts of clouds while B6 shows differences in soil mineral content and could indicate ice.
By comparing the four winter charts it can be seen that November of 2014 was definitely the coldest and toughest one for the birds. Only one month later in December it got a bit warmer and ice started to melt. Moreover, clouds on the sky decreased. In comparison to the winter months from the previous year the winter of 2015 was definitely warmer as it can be seen through the bar charts. There were fewer clouds, ice and snow. It can be said that this was probably the most bearable winter for the remaining birds which stayed in Mistford and did not fly to the South. However, one major indicator for these rapid and fluctuating weather changes is definitely the global warming which has a big impacts on the weather and the environment and will continue to effect the future.

Summer
During the summer months June, August and September floods and burned land could be detected through the multi-spectral values B2, B4 and B5. B2 indicates different plant types while B5 is completely absorbed by liquid water and is sensitive to moisture content of soil and vegetation. Moreover, B4 is partly absorbed by liquid water and is sensitive to vegetation structure and chlorophyll. Comparing all the charts from these three years it can be seen, that B4, which is sensitive to the vegetation structure and to chlorophyll has always been the highest values among these measurements. Combining the high chlorophyll and the varying B2 amounts it can be speculated that all of the summer months must have been pretty warm and the plants could operate photosynthesis. Furthermore, it can be seen that June of 2015 and 2016 have been the hottest months during the measuring period due to the high B4 values. As the B4 value has been always higher then the B2 value it could also indicate that it was too hot sometimes for the plants and land got burned due to the weather conditions, which has led to a decrease of the different plant types and an increase of burned land. Therefore, that June has been probably the hottest month, it can bee said that the decline of plant types has erupted especially in this month. Nonetheless, also in September, 2016 a massive decrease of plants can be noted in comparison to August, 2014.
Another condition, such as floods might have occurred during these summer months due to the B5 values which indicate clouds, water and moisture. The highest B5 value can be found in the chart of June, 2015. Having a closer look to all the different B-values of this view it can be said, that the summer of 2015 must have been the toughest one, due to the hot phases which have lead to burned land. Due to the flooding, the ground might have gotten exhausted. Furthermore, it can be extracted that June, 2016 has been quite strenuous as well as the highest amount of burned land can be recorded during this period. In comparison to June from the previous year there have been fewer floods.
To sum up it can be said that the summer months August, June and September of the last three years have been quite stressful for the habitat which has lead to some environmental changes that are affecting the bird's population as well.

Comparison of the NDVI Values
In order to compare the habitat's surface we visualized our data and the NDVI values through heat-maps, so that we can draw conclusions about the changes. While focusing on the weather conditions and the health of the plants in the previous charts we wanted to get more detailed information about the ground through these heat-maps. Generally it can be said, that the surface of the nature habitat has changed according to the seasons, which means that we could observe that during the winter months a lot of snow has fall which might have frozen to ice due to the overbalanced low NDVI values. A low NDVI is an indicator for a rarely grown surface, while a high NDVI states a flourishing vegetation. As it got warmer in February and March it can be seen that snow and ice were slowly melting and the NDVI grew. In comparison to the NDVI values from the "Winteränd Late Wintercharts, the SSummer"views include lot higher NDVI values. Furthermore, it can be seen that the plants look healthier and that the vegetation in the summer months is flourishing again.
After having a closer look at the changing surface of the preserve, we could detect one major transformation. Examining especially the summer months we have seen that the shape of Lake Boonsong has changed over the months. We belief that one of the companies is pouring chemicals into the lake which leads to a pollution of the lake and to a drastic change of the surrounded environment. We see a massive endangerment of Lake Boonsong, as for the birds living within the preserve and for humans as well! The chemicals which have been poured into the lake might vaporize and arise into the air which would lead to an even worrying air pollution which would affect in first place the birds living in the preserve but also visitors and rangers working and entering the nature park. Furthermore, the poisonous water would have alarming impacts on the surface's quality. Not to mention what effects it could have on the people living nearby the preserve when the polluted water would get into the groundwater cycle.
We are concerned about the Blue-Pipit's living space, if nothing changes about the current state of Lake Boonsong. Post-doc student Mitch needs to pursue further steps in order to safe the nature park's wildlife and to prevent further environmental damages.

3.4 Grand Challenge

3.4.1 Description of the Dashboard

This visualization represents the Grand Challenge, for this view we decided to take the most important views from the previous Challenges and combined them into one dashboard, which will give a good overview of the preserve's situation. Since we have changed a lot within the single Challenges, the Grand Challenge has changed as well. The first row includes three heat-maps which contain the NDVI values from three different years. We took from each year almost the same month of data to provide the most accurate comparison.
Instead of the more general view as we have planned at the beginning, we added right under the heat-maps the most important view from Challenge 1, which states the car-types that drove at a certain time through the gates. Due to this chart it is possible to examine when and what kind of car drove through the park, which gives more clearance about the usage of the gates per day. Here the user can filter the car-types, gate-names and hour to take a closer look at the different cars passing the preserve.
Last but not least the last row of this dashboard presents the second Challenge. The left view shows a bar chart which states the different chemicals that have been detected in the air. Moreover, the view on the right side demonstrates the activity of the sensors over three months. Additionally, these two charts interact with each other. If a user selects one of the lines from the Äctivity of Monitors", the chemical-view will automatically change according to the chosen data and will show which chemical has been detected through the chosen sensor.
Furthermore, we added a line graph which we created though Excel, which visualizes the textual data from the Newsletters of Mistford. In order to apply these data we copied the information from the Newsletters into an Excel Sheet and draw a graph to visualize the financial situation of each company in Mistford.

3.4.2 Reasons for our Design Choices

As we already discussed the single visualizations of this dashboard in the previous chapters, we decided to focus in this section why we arranged the different graphs for the Grand Challenge the way we did. To build this dashboard, we took the most important views from each Challenge and combined them into one view.
Firstly, we put the NDVI-value comparison on top of the dashboard in order to give the user an overview of the health condition of the plants within the preserve as well as about the environmental changes that have been taken place between Summer 2014, Fall 2015 and Fall 2016. Furthermore, we also provided a legend which makes it even easier to interpret the colors and the meaning of this views.
Right under these graphs we added the most important view from Challenge 1 where it can been see which car drove through which gate at what time. Here we provided a filter to search for specific details as for the gate-names, hours and car-types. Moreover, we used different colors for a clearer structure and better overview.
The two graphs in the middle of this dashboard present the data from Challenge 2. Therefore, we added the bar chart which labels the measured chemicals in the area of Mistford. The other view shows the operation activity of the monitors and makes it possible to examine if the sensors were working faultlessly or not. Moreover, we implemented filters and tooltips to analyze the data even more interactively.
Furthermore, we also added the chart which we created via Excel which states the financial status over time from the different companies. In order to be able to see the exact course from each company we decided to user different colors.

3.4.3 Results

The results from the Grand Challenge were, according to the data of Challenge 1 that there is suspicious and forbidden traffic within the areas from ranger-stop 3 to gate 3 and between gate 5, ranger-stop 6 and gate 6. Moreover, we noticed that the sensors measured cars passing these gates during night time. Furthermore, we examine that cat-type 4, a vehicle with three or even more axles, is responsible for causing traffic. This could mean nothing, but it is still suspicious. Even though the cars are simply passing through the park, they still interfere with the bird's habit. However, which could be a more likely reason for these cars driving though the park at night, might be some illegal issues. Having a closer look at the newsletters might give some clearance from which company the trucks are coming. For example, a transport employee from Kasios Office Furniture said in an interview that he has to work long and late hours.
After examining the data from the newsletters we could see that there was a heavy decline of finances at Kasios and Radiance in December 2014 which was caused by stricter environmental regulations. However, it is pretty suspicious that the Kasios Office Furniture company experienced an increase of their financial situation right away from this date. And their financial aspect in August 2016 is even better than it was before the decline in 2014. The company must have found a way to compensate the heavy losses very quickly.
Since we already found out that it is very likely that Kasios Office Furniture caused illegal traffic at night, this could be one of the reasons why their income has increased again. Thinking of any illegal subjects the company could do in the middle of the night with their trucks in the preserve, illegal trash dumping could be one possible reason, why the company's income increased that quickly again.

4 IMPLEMENTATION

4.1 Toolkit & Platforms

We decided to use Tableau as the main tool for this project, because we already used the it for previous tasks during the VIS course. As the Challenges are about finding certain reasons for environmental changes we believe that this tool is very suitable to accomplish the given tasks.
Tableau includes a wide range of tools, such as the option to add various filters, brushes or highlights, which we used for our implementations. For handing in our visualizations we used Public Tableau where we uploaded the dashboards of the four Challenges and provided access trough URL links to our work.
For the Grand Challenge we decided to include one view which we created by using Excel in order to visualize the data from the newsletters we received. Therefore, we took the data from the PDF files and generated a line graph through the extracted data.

4.2 Challenges during the Implementation Process

The main problems we were facing during the implementation process were developing the right visualizations for the tasks we had to accomplish. It took us a while to find the suiting view for the NDVI presentation for Challenge 3. In the end we decided that a heat-map would be the best solution for this task. Moreover, loading the dashboards into Public Tableau was a bit challenging as well, because we had to handle a lot of data. Therefore, we tried to reduce the data as much as possible and created extracts.

5 DISCUSSION

Even though, some computer scientists might say never change a running system"we do believe that there are always some improvements that could be adjusted to improve visualizations to make them more user-friendly. Furthermore, we also think that it is important to reflect one's own work after finishing and to observe it critically by trying to identify its strengths and weaknesses.

5.1 Strengths and weaknesses

Even though, we had predefined tasks from the VAST community which we should accomplish and achieve in the end of our project, we did not only put the focus on these tasks, but also tried to work with the data individually and experiment with different data types, diagrams and visualization techniques, hoping to find out something exciting and special about the data which was not necessarily requested. However, we tried to combine our imaginations of what was important to us and the tasks which needed to be done for this project in order to hand in a successful assignment.

However, we tried to arrange the visualizations of the various Challenges appealing to the user's eye, some might criticizes the layout of the dashboards. The more graphs a dashboard includes the smaller they get. It was quite a challenge to structure our dashboards in an effective an still readable way, due to the different diagram types and their contents. Another aspect which could be argued about, are our color choices. We think that it is sometimes impossible to choose a fitting color design which will be approved by everyone. Some people might think that we have used too much colors instead of working with different hues. However, we put a lot of thoughts into our color concept in order to be able to express the information we want to focus on. Although we have learned during our VIS classes that using different hues is indeed a better option to differentiate, but we do think that the colors we have used for our project will not overburden the users nor distract them.

In addition to the colors, we also implemented interactions to illustrate connections between the diagrams. While the first three Challenges do profit from their implemented filters and other interactions, Challenge four is rather poorly equipped when it comes to interactions. It does offer tooltips and highlighting specific areas when chosen, but the single graphs within the dashboard do not interact. As the dashboard of Challenge 4 is intended to give an overview of the previous three Challenges we were not able to find any connections between the different data types. Nonetheless, to us it was very important to create a system which is easy to use and to understand without detailed instructions. Moreover, we also used simple and self-explaining headings for each graph to ensure, that the user knows only by having a quick look at the dashboard what it stands for and what kind of information it contains.

Having taken all the positive and negative aspects of our work into account with, we can say that we are pleased with the work we have done. Nonetheless, nobody is perfect and we are sure that there are more aspects then now mentioned which could be edited or even redefined.

5.2 Lessons learned

One of the most important skills in our opinion which we have gained during this semester, while working on our group project is the right way of how to use the visualization tool Tableau. While a lot of people might think that Tableau is quite easy to use due to its drag and drop system, it can be tricky and complicated to operate with, especially when trying to demonstrate a certain part of the given data. The complex part of using Tableau is the right way of data preparation. With Tableau you have a lot of opportunities to represent data through various data types, which means that you really have to be precisely. Some students might underestimate this tool, but after getting a closer insight into the different functionalities of Tableau we do believe that it can be very helpful to organize huge datasets and to present them in an understandable way.

Especially, in today's world where the amount of information is increasing enormously this project was a very good practice to understand how to handle different data types and how to extract the most important information out of it, in order to present it the most fitting way. Moreover, we are sure that we will be using this program in the future, either for university projects or later on during our professional life.
Furthermore, we have understood how much LO-FI prototypes might vary from the HI-FI products. At the beginning we had a brief look at the data which we received, and started to design our low-fidelity prototypes. As we finished our hand-drawn prototypes we had the feeling that we have done a quite good job and had the impression that we could realize these ideas through our implementation tool. Admittedly, we acknowledged quickly that the concept which we had created earlier was not as good as we thought. Through this project we have experienced how important it is to create low-fidelity prototypes in order to have a rough idea about how to develop the project which needs be refined and improved during the whole implementation process.

However, choosing a fitting way to present data is one important aspect which needs to be considered while developing visualizations. On the other hand creating interactions and relations between the single graphs is another very significant characteristic which needs to be thought of. By experimenting with different visualization techniques such as filters and tooltips we have understood how essential they are in order to express relations between diagrams or to emphasize the meaning of them.

To sum up all the skills we have gained and learned during the implementation of this project it can be said that we really enjoyed working with Tableau even if it was quite tricky at the beginning.

6 Separation of Tasks

We tried our best to divide the tasks as fair as possible between us two. Therefore, we decided to split the required tasks for this milestone and final report the following way:

Bianca Barnucz

- Updating the Website
- Abstract
- Motivation
- Related Work
- Approach (Challenge 2 & Challenge 3)
- Results of Challenge 2 & Challenge 3
- Discussion
- Adapting the Tableau implementation of Challenge 2

Stephanie Wegscheidl

- Updating the Website
- Abstract
- Motivation
- Related Work
- Approach (Challenge 1 & Grand Challenge)
- Results of Challenge 1 & Grand Challenge
- Implementation
- Adapting the Tableau implementation of Challenge 1

7 References

[1] https://www.youtube.com/watch?v=c3DY_hK7NhQ&t=103s

[2] https://www.gicentre.net/vastchallenge2017/

[3] https://vimeo.com/225778875

[4] "VAST Challenge 2017 MC3 Solutions Final"

www.cs.umd.edu/hcil/varepository/VAST%20Challenge%202017/challenges/Mini-Challenge%203/solution/VAST%20Challenge%202017%20MC3%20Solution%20Final.pdf

8 Public Tableau Links

- https://public.tableau.com/views/1_Challenge-FIN/Challenge1?:embed=y&:display_count=yes&publish=yes
- https://public.tableau.com/views/M4-Challenge2/Dashboard1?:embed=y&:display_count=yes
- https://public.tableau.com/views/M3_Challenge3/EnvironmentalChanges?:embed=y&:display_count=yes
- https://public.tableau.com/views/Grand_Challenge/FINALDASHBOARD?:embed=y&:display_count=yes&publish=yes