SMART MONITORING SYSTEM FOR BEEKING FARMS

A PROJECT REPORT

Submitted by

Rahul Sanjay Mandviya (19MIM10062)
Abhinav Gadgil (19MIM10067)
Yash Sharma (19MIM10068)
Shivam Bajpeyi (19MIM10082)

in partial fulfillment for the award of the degree of

INTEGRATED MASTERS OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

specialization in

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
VIT BHOPAL UNIVERSITY
KOTHRIKALAN, SEHORE
MADHYA PRADESH - 466114

OCTOBER 2020

Ref: Hex-Community/Interns/VIT-Bhopal/BeeHives/2020/01 **Date:** 22/10/2020

TO WHOM IT MAY CONCERN

This is to certify that **Below Mentioned Students,** has successfully completed all tasks assigned to them as Part of our **Phase-1 of Beehives Monitoring Project using Statistics & Predictive Modelling Techniques** for Beep Keeping Farmers in Maharashtra, India. During the period of this **Internship Program** with us, all of them were found Punctual, Hardworking and Inquisitive.

1. **Rahul Sanjay Mandviya**
2. **Abhinav Gadgil**
3. **Shivam Bajpeyi**
4. **Yash Sharma**

As we evolve our BeeHives Monitoring Product on-going basis, we would be happy to engage all above students in our Data Science Research, Consulting Projects and AI/ML Community Development Programme in the field of Data Science.

We wish them every success in life.

For Hexaberry Technologies LLP

MANGESH CHINCHOLKAR
Co-Founder & Partner

HEXABERRY TECHNOLOGIES LLP
Office Address: Evoma Business Centre, Prestige Featherlite Tech Park,
Plot No 148, 2nd Phase, EPIP Zone, Whitefield, Bangalore – 560 066, India
Phone: +91 - 80 4903 3082, **Mobile:** +91 - 96 8675 4340
Registered Address: Sraddha Gardenia, B-102, Rahath Bagh, C.V. Raman Nagar, Bangalore – 560093
(LLPIN: AAQ-1909, PAN: AAMFH4144A, TAN: BLRH10871A, GSTIN: 29AAMFH4144A1ZK)

www.hexaberrytechnologies.com

VIT BHOPAL UNIVERSITY,KOTHRIKALAN, SEHORE MADHYA PRADESH – 466114

BONAFIDE CERTIFICATE

Certified that this project report titled **"SMART MONITORING SYSTEM FOR BEEKEEPING FARMS"** is the bonafide work of "**RAHUL SANJAY MANDVIYA (19MIM10062), ABHINAV GADGIL (19MIM10067), YASH SHARMA (19MIM10068), SHIVAM BAJPEYI (19MIM10082)"** who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported here does not form part of any other project / research work on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

PROGRAM CHAIR
Dr. Pandimurugan Vellaisamy,
Assistant Professor
School of AI &ML division
VIT BHOPAL UNIVERSITY

PROJECT GUIDE
Dr. M. Ashwin,
Senior Assistant Professor
School of AI &ML division
VIT BHOPAL UNIVERSITY

The Project Exhibition I Examination is held on 27th October 2020

ACKNOWLEDGEMENT

First and foremost, I would like to thank the Lord Almighty for His presence and immense blessings throughout the project work.

I wish to express my heartfelt gratitude to Dr. Nageswara Guptha M, Head of the Department, School of Computer Science and Engineering for much of his valuable support encouragement in carrying out this work.

I would like to thank my internal guide Dr. M. Ashwin, for continually guiding and actively participating in my project, giving valuable suggestions to complete the project work.

I would like to thank all the technical and teaching staff of the School of Computer Science and Engineering, who extended directly or indirectly all support.

Last, but not the least, I am deeply indebted to my parents who have been the greatest support while I worked day and night for the project to make it a success.

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
1	Beehive	11
2	Beehive Frames	11
3	Work Breakdown Structure	12
4	Figma Wireframe	13
5	Univariate Graphs	16
6	Time-series Graphs	17
7	Scatter Plots	19
8	Line Graph	20
9	Add Location Screen	21
10	Home Screen	21
11	Glimpse of our Project	23

ABSTRACT

The main objective of our project is to improvise the situation of beekeepers by implementing data analysis based on real-time data. We initiated our project by overviewing some research articles to get a theoretical viewpoint of our topic. Once the research was complete, we collected raw data from actual farming sites. Upon collecting data for at least two-three months, we visualized the graphs to breach some problems such as extreme values, poor quality data, inaccuracy in values, inaccessible data. After finding solutions and making the data more "useful", we implemented in-depth data analysis using python libraries such as pandas and NumPy. Thereafter we created time-series graphs to identify the pattern that is followed on a daily basis. We also plotted graphs between two variables from the dataset to get a better idea of how the variables are correlated. We even implemented descriptive statistics to calculate the important statistical values and check the quality of our datasets. Once our analysis was complete, we developed an application for the welfare of beekeepers. All the beekeepers We characterized our application with a simple interface so that any new beekeeper who is not acquainted with technology can easily access it. The application basically displays the location of the farm and number of beehives and the beekeeper has the liberty to add his farm location and the respective number of beehives residing in that location. We even added a static graph to display some rough insights.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	List of Figures	**5**
	Abstract	6
1	**INTRODUCTION**	
	1.1 Introduction	8
	1.2 Motivation for the work	8
	1.3 Our Idea for project	8
	1.4 Problem Statement	8
	1.5 Objective of the work	9
	1.6 Organization of the thesis	9
2	**LITERATURE SURVEY**	10
3	**PROJECT PROCEDURE**	12
4	**WORK DONE**	
	4.1 Figma	13
	4.2 Data Analysis	14
	4.3 Flutter	14
5	**OBSERVATIONS**	16
6	**RESULT & CONCLUSION**	22
7	**RECOMMENDATION FOR FUTURE WORK**	31
8	**REFERENCES**	32

1. INTRODUCTION

1.1 Introduction

Beekeeping (or apiculture) is the maintenance of bee colonies, commonly in man-made hives, by humans. Most such bees are honeybees in the genus Apis, but other honey-producing bees such as Melipona stingless bees are also kept. A beekeeper (or apiarist) keeps bees in order to collect their honey and other products that the hive produces (including beeswax, propolis, flower pollen, bee pollen, and royal jelly), to pollinate crops, or to produce bees for sale to other beekeepers. A location where bees are kept is called an apiary or “bee yard”.

1.2 Motivation of the work

Hexaberry Technologies LLP engaged with Bee Keeping Farmers in Maharashtra (Shegaon & Harali) to collect Temperature, Humidity and Weight data of Beehives for effective Monitoring in order to take informed decisions to improve upon revenue from honey yield that gets collected at each of the Beehives box. Honey contains a wide range of sugars, varying according to the nectar source, and small amounts of other substances such as minerals, vitamins, proteins, and amino acids. Everyone knows that honey is extracted from honeycomb. But have you ever given a thought that the traditional method of burning honeycomb in order to put the bees away sometimes kills a lot of bees which results in less production of honey in the future due to the reduction of bees in the environment. So, we decided to create such an environment for honeybees which helps them to survive for a long time that’s an aid to increase honey production.

1.3 Our Idea for project

We got to know that farmers in India have also started using the modern techniques of beekeeping but they still rely on their own intuition for deciding whether a wet gunny bag should be kept on the beehive or not. This leads to human error. So we decided to collect data of temperature, humidity and weight of the beehive in order to perform some analytics. This analytics could help us to understand when the wet gunny bag has to be kept on beehive, how often the beehive has to be

emptied. Once we have this information we can have a simple mobile application which can help farmers make correct decisions easily.

1.4 Problem Statement

Can we develop a smart solution that can monitor weather conditions in the beehive and can tell farmers, how to control weather conditions in the beehive and take necessary steps in order to get the best yield through a simple app?

1.5 Objective of Work

Being a AI student, we thought why not do a detailed analysis on the data collected from the beehive and understand how the whole process works. Once we get some insights through analysis, we can develop a simple, easy to use mobile application which will help farmers to get to know what are the necessary steps to be done.

1.6 Organization of the thesis

Chapter 1: Introduction
Chapter 2: Literature Survey
Chapter 3: Project Procedure
Chapter 4: Work done
Chapter 5: Observation
Chapter 6: Result & Conclusion
Chapter 7: Recommendation for future work
Chapter 8: References

2. LITERATURE SURVEY

Beekeeping is influenced by many external and internal factors that determine the lifecycle of bees. Bees are an important part of our ecosystem, as they prove to be top contributors for pollination and other bee products as well. The quality of agricultural production directly depends on pollination. When the queen bee lays eggs at that point of time the temperature in the beehive rises up to 42 degrees Celsius and then the temperature of the beehive is reduced by worker bees by ventilating the beehive through their wings and by evaporating water content in honey, thereby making honey viscous. The yield from beehive farms could be halved if the temperature exceeds its threshold value. Now during summer as the outside temperature is very high, bees spend loads of time and energy to cool their beehives, thereby reducing their navigation radius. If the bees fail to maintain internal temperature, the comb starts melting resulting in their demise. Due to this, the yield of all the other products like bee wax, propolis pollen, etc. also reduces. It is also observed that during this time beehives are prone to moth hatchlings and parasite attacks. They consume their offspring and devastate entire colonies affecting their structural patterns. In India, the beekeepers control such outcomes based on their traditional knowledge, experience, and observation skills. Hence it is difficult for them to take precautionary measures against such unpredictable mishaps. They put wet gummy bags on the beehives in an orderly schedule, based on their intuition. If they feel that outside weather is hot, then they put the wet gunny bag on the beehive but there might be a case wherein the actual temperature in the beehive might not be above the threshold value. The bee population ratio is also an important factor that determines the yield of bee products. Since most beekeepers are concerned with the honey, they often check the weight of the box. However, it is important to consider that lifting these boxes disturbs the bee formation and it may affect their process. These all inconsistencies and various observations contribute to inefficiencies in beekeeping, making it a completely different subject to study. The fig[1] shows how a beehive looks and the fig[2] shows frames in beehive.

Fig[1]: <u>Beehive</u>

Fig[2]: <u>Beehive frames</u>

3. PROJECT PROCEDURE

As our project had lots of tasks to be accomplished, we planned to use the Agile Development methodology for our project. First, we listed all the tasks that are to be done in the project. Then we added them into sheets in a sequential manner in order to make a Work Break Down Structure. Then we added task dependencies. After making the task list we divided ourselves in a group of two people each for every task. In this one group had to complete the task assigned to them and the other team had to test and review the work done by the team. Once the completion and testing of the task was done it was marked as completed. This helped us to keep track of the amount of project done until now and the amount project to be done in the coming days. This helped us to work in an organized manner and helped us to complete the tasks in the given time. Fig[3] shows some glimpse of our work breakdown structure.

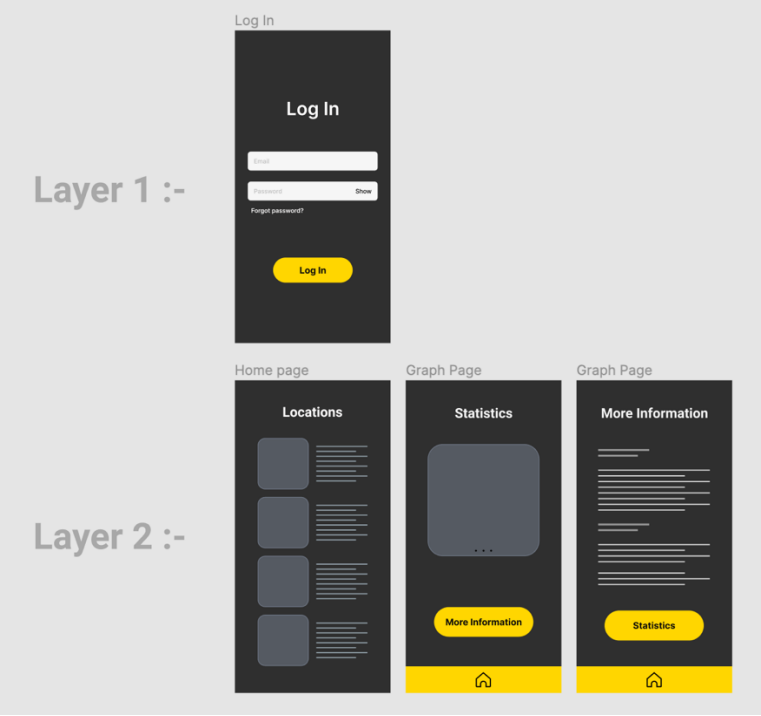
Sr. No.	Task Name	Task Description	Technology Outlook	Technologies Required	Comment	Task Dependency	Code File	Assigned to	Reviewed by	Status
Mobile APP										
1	App Home Screen	Information of product. New User, Existing User buttons.	Mobile App	Flutter and dart	Flutter for framework and Dart for programming	None	Startpage.dart	Rahul, Shivam	Abhinav, Yash	Completed
2	Enquiry Screen	New user button clicked - give contact details	Mobile App	Flutter and dart	Flutter for framework and Dart for programming	App Home Screen		Abhinav, Yash	Rahul, Shivam	Completed
3	Log In Screen	Existing user button clicked - display log in option	Mobile App	Flutter, dart, Firebase	Flutter for framework, Dart for programming and firebase for backend	App Home Screen	Login.dart	Rahul, Abhinav	Yash, Shivam	Completed
4	Location screen/Home Screen for existing User	sites, summary of beehives, sites will be in button format - when clicked will take to graph screen, Support/help Button	Mobile App	Flutter, dart, Firebase	Flutter for framework, Dart for programming and firebase for backend	Database knowledge, LogIn screen	Home.dart	Abhinav, Yash	Rahul, Shivam	Completed
5	Analytics Screen	Graphs of temparature, humidity, icon for weight, Button for alerts	Mobile App	Flutter, dart, Firebase	Flutter for framework, Dart for programming and firebase for backend	Data Analysis		Yash, Shivam	Rahul, Abhinav	Completed
Data Analysis										
1	Dummy Dataset	Create raw datasets of Temperature, Humidity, Weight	Dataset	Excel	Create raw datasets to test our analysis before using actual datasets			Yash, Shivam	Rahul, Abhinav	Completed

Fig[3]: Work Breakdown Structure

4. WORK DONE

4.1 Figma

As we know that farmers of India are not more technically sound, we had to make our app farmer-centric and give it a very simple easy to use interface. We brainstormed and kept ourselves in farmer's shoe in order to decide what features our app have. We had to keep very minimum features in app so that farmers don't get confused. For this we talked with some beekeepers and showed them the interface designs and came up with the wireframe shown in the fig[4]. We used Figma for wireframe designing because of its simple drag and drop functionality and collaboration features. Moreover, figma's education feature allows unlimited people to collaborate and work together for designing wireframe.



Fig[4]: Figma Wireframe

4.2 Data Analysis

We used python libraries like Pandas, Matplotlib, and Plotly, and Jupyter Notebook for data analysis. The very first challenge in front of us was to prepare clean data from the raw data that we collected from the farms. Our data had mainly three features - Temperature, Weight, and Humidity. Using the Pandas library, we first calculated the number of missing values in each feature as there might be certain events wherein the sensors didn't collect any data. After calculating the number of missing values, we deleted the empty rows in order to get clean data. This new clean data was stored in the ".CSV" extension file. Now the new data set was loaded, and the frequency distribution of each feature was studied in detail. Further to get a complete understanding of data, we visualized data using univariate graphs i.e. histogram through python library matplotlib for each feature. After this, we performed some inferential statistics in order to get an in-depth understanding of data. We further plotted time-series graphs which helped us to understand how the features vary with time and gave us some meaningful insights. Now, in order to understand the relation between the features, a correlation mat was plotted, and bivariate graphs i.e. scatter plots were used. In this project, we need to take complete care of every minute detail which we get from visualization. So, for time-series graphs and scatterplots, we used plotly instead of matplotlib for visualization. Plotly helps us to plot interactive graphs wherein we zoom in and observe each and every minute detail of data. Moreover, when we hover our mouse onto a certain region of the graph it shows the values of the coordinates.

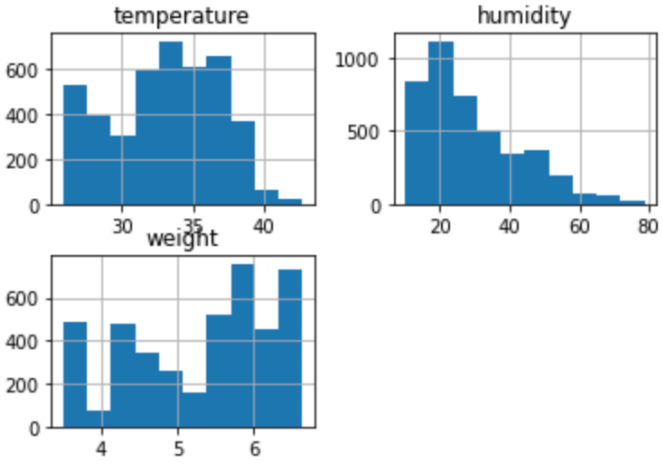
4.3 Flutter

To showcase our analysis to the farmers in an efficient way such that it helps them to maximize bee production. We have decided to make a user-friendly android app using Flutter. Basically, Flutter is an open-source UI software development kit created by Google. It is used to develop applications for Android, iOS, Linux, Mac, Windows, Google Fuchsia, and the web from a single codebase. For coding, we have used a dart programming language that is supported by flutter. Basically, Dart is an object-oriented, class-based, garbage-collected language with C-style syntax. Dart can compile to either native code or JavaScript. It is developed by Google and is used to build mobile, desktop, server, and web applications. First, we developed the first screen of our app where we had to show the logo, name, and motto of our project. Then, for giving access to our app to the users we have created a Sign-Up screen for new users and a Login screen for existing users. For easier navigation,

we added a bottom navigation bar so that users can access the actual functionality of the application. After the login, we have created the My Location screen where we have given the option of adding new locations of bee boxes to the users for monitoring. This screen shows the name of the location and the number of bee boxes at a particular location. Whenever the user clicks on the name of the location then the next screen of our app will come up i.e. Graphical Representation screen. In this screen, we have shown our analysis through graphs. Since we have not included IoT devices in our project till now, we have plotted a static graph. In this screen, we have also added the basic information of beehive temperature, humidity, and weight and at last, we have created an FAQ screen where we added the general questions and their answers that help beekeepers. Excluding the mentioned screens, we have also created some widgets like logout, home screen, and multiple options widget on each screen just after the login screen.

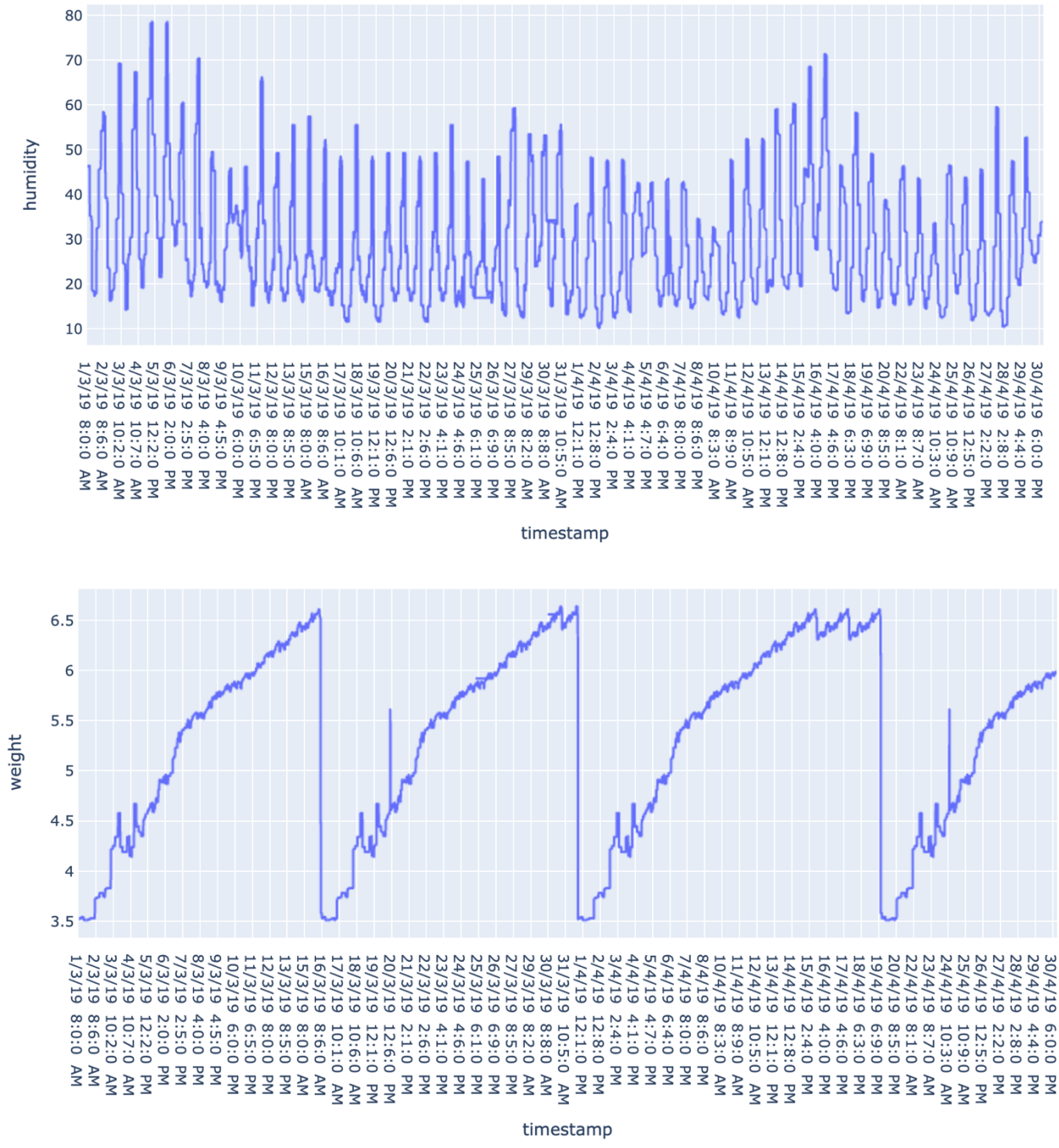
5. OBSERVATIONS

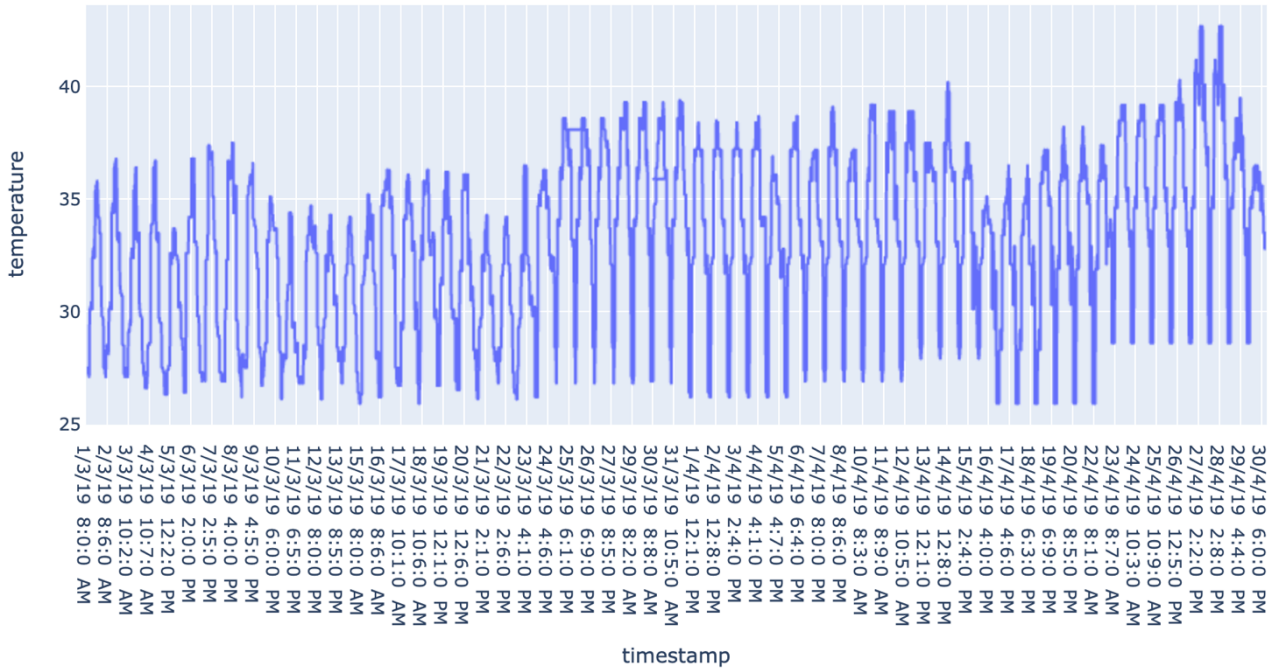
It was observed that there were few missing values in the raw data. After removing the empty rows, inferential statistics was performed on to the data which gave us some important information like minimum temperature, humidity, and weight of the beehive was 25.9 degrees Celsius, 10.1% and 3.51 kg respectively and maximum temperature, humidity and weight of the beehive was 42.7 degrees Celsius, 78.6% and 6.64 kg respectively. Later the univariate plots shown in fig[5] gave more information about the distribution of data. We observed that humidity data is left-skewed or exponential and temperature and weight are almost normally distributed.



Fig[5]: <u>Univariate Graphs</u>

On observing time-series graph we got to know - humidity is high in the morning. It reduces significantly in afternoon at around 2pm. And then increases in the evening. In the month of March, the frames in the beehive were emptied twice, once on 16/3/19 and second time on 31/3/19. But in the month of April it was emptied on 20/4/19. So, from this we could infer that yield is higher in month of march than in April. Temperature of beehive rises up to a great extent in time period 12-4pm. It goes above threshold value of 32-33 degrees Celsius which can affect bees a lot and can result in a bad yield. Fig[6] shows time-series plot for humidity, weight and temperature.

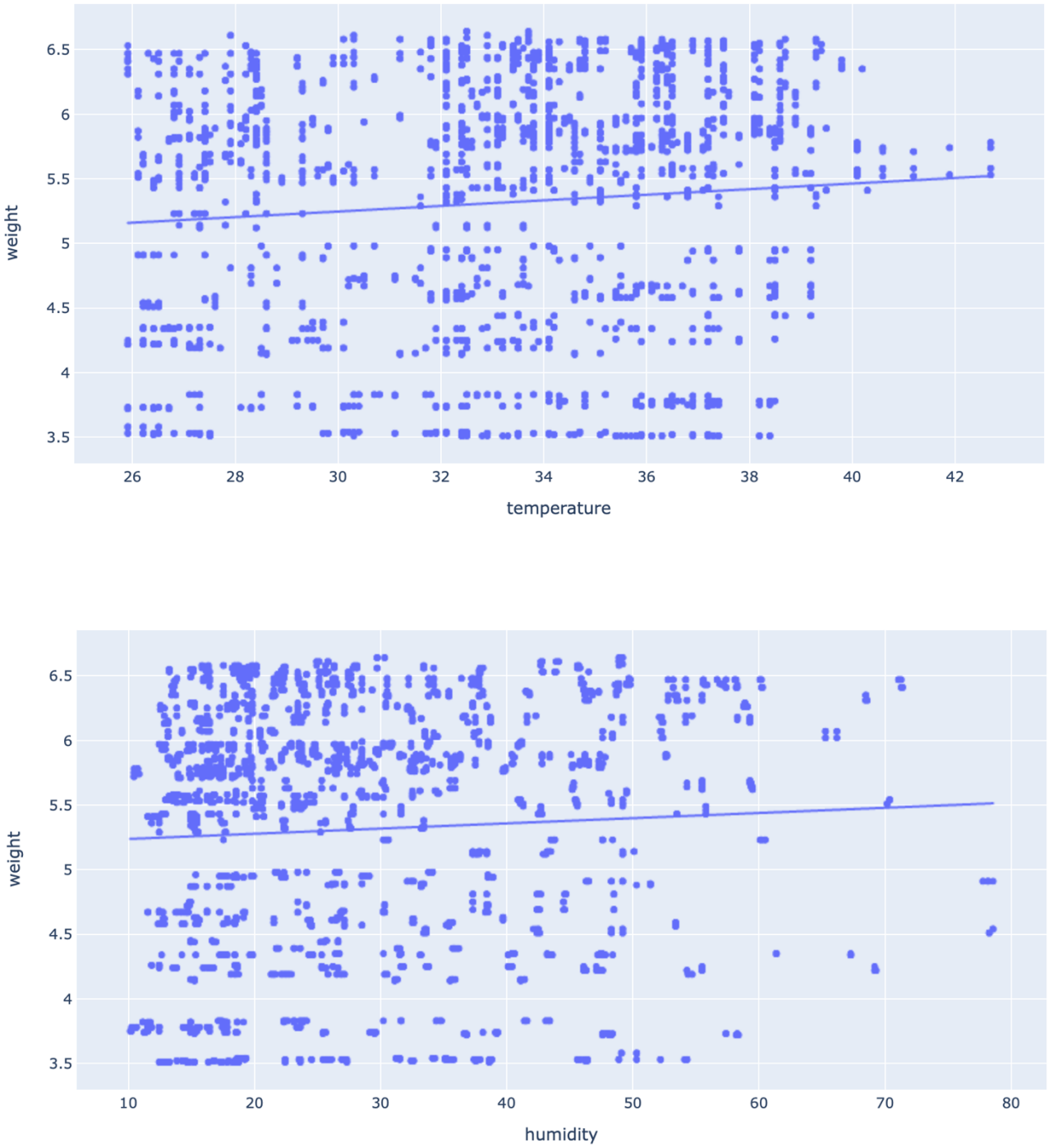




Fig[6]: <u>Time-series Graphs</u>

Later, we calculated Pearson's correlation between the features in order to understand how the features are affecting each other. Humidity and Temperature are inversely proportional having a strong negative correlation of -0.7140. There is a weak to moderate correlation between humidity and weight and also between temperature and weight. The correlation between humidity and weight was calculated as 0.05784 and correlation between temperature and weight was calculated as 0.08703.

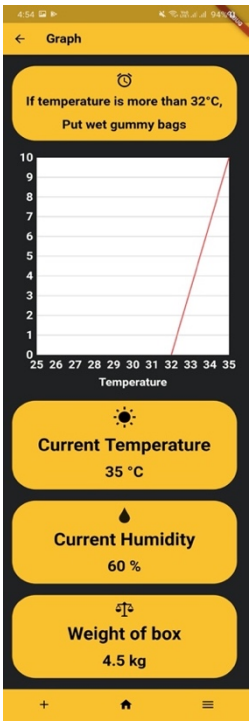
Further, scatter plots given in the fig[7] gave us detailed information that in afternoon when temperature is very high and humidity is very low the weight in the beehive rises. Through this we can infer that if the queen bee lays eggs in afternoon time then the worker bees have to spend a lot of time in ventilating the beehive as outside temperature is also very high. Moreover, bees instead of going out to collect nectar keep ventilating the beehive which results in less collection of nectar and reduces the overall yield. If the temperature is not lowered, then this can also lead to melting of the beewax.



Fig[7]: <u>Scatter Plots</u>

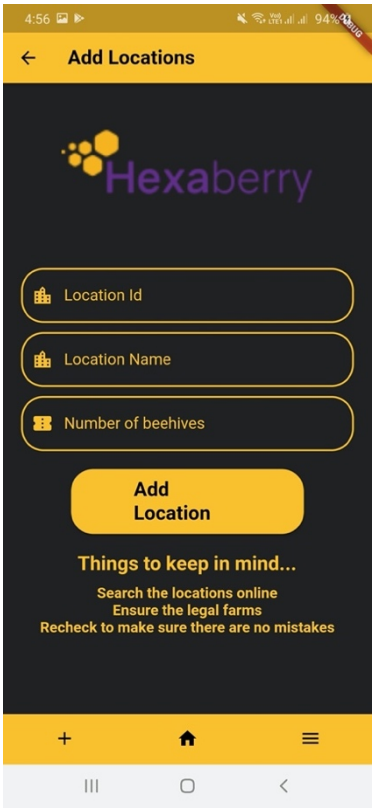
Further we showed few graphs to farmers and got to which graphs are very easily interpreted by the farmers. Line graph helped farmers to simply understand that when line is inverted upwards, he has

to put wet gunny bag on the beehive and when it is inverted downwards then, he should remove the wet gunny bag. Fig[8] shows one such graph.

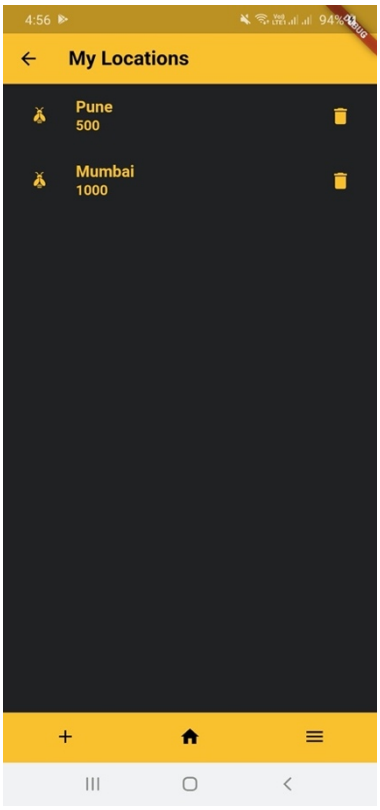


Fig[8]: <u>Line Graph</u>

While testing our app we observed that there might be a case wherein a farmer might have two beekeeping farms in vicinity. In such cases the database might crash. For this we added Location ID feature in our add location screen in fig[9] and we added one simple home screen interface shown in fig[10], where they get info about how many beehives are there at each location they have added.



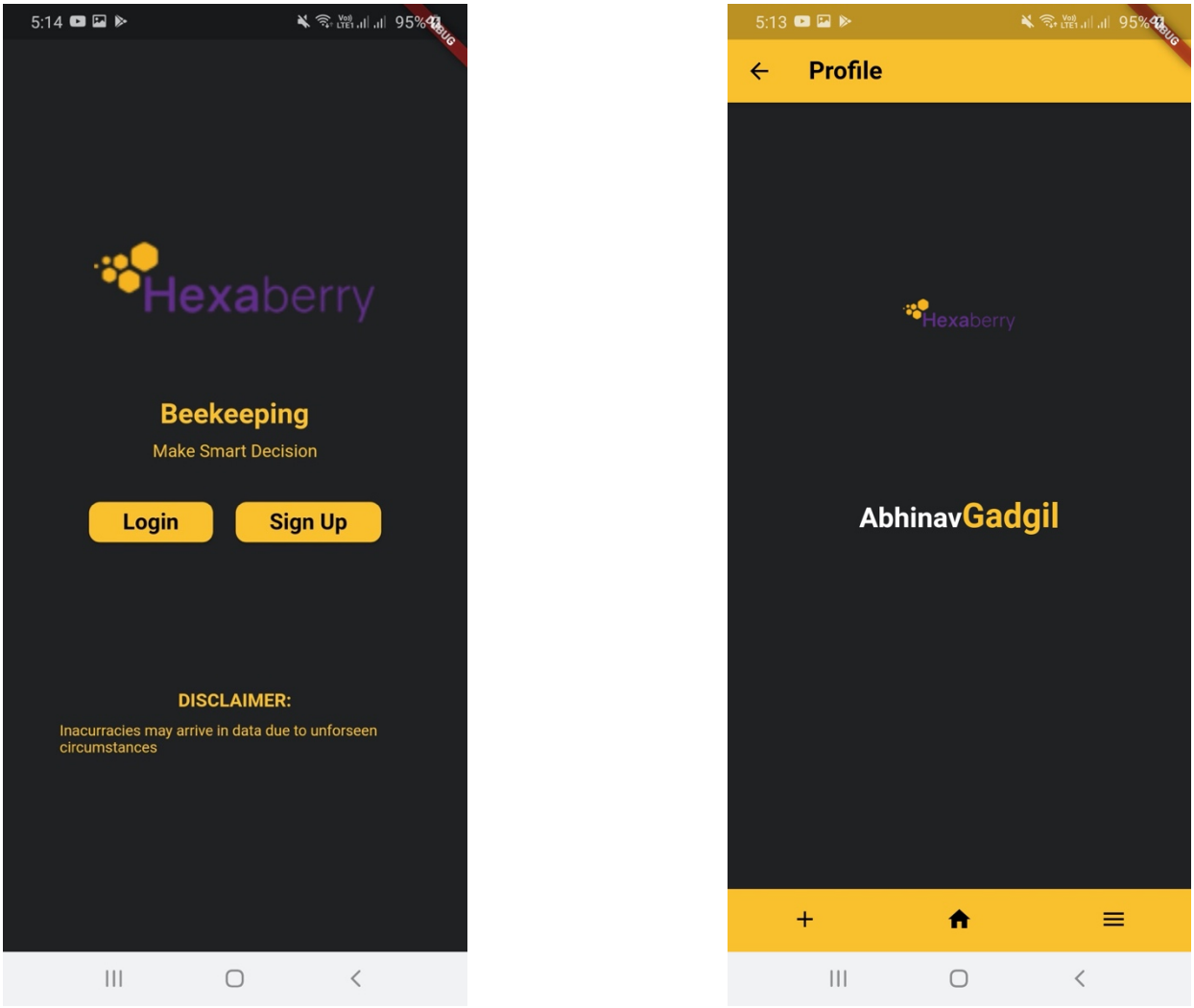
Fig[9]: Add Location Screen



Fig[10]: Home Screen

6. RESULT AND CONCLUSION

Inferentially, the project managed to give us hands-on experience with actual data and app development. After overviewing the observations, we can conclude that analysis of beekeeping is filled with intrinsic complexities that require tons of experience handling data. When implementing data visualization, we discovered many adverse and irreparable issues by which we were trapped on a single phase for a long duration. Though problems were breached at the end we got acquainted with the apiculture and the depth of how each factor is correlated. Most importantly our work also gave us broader aspect of “what is important” to get useful insights. After comparing our expected and observed results, we came to a conclusion that it is essential to include more variables in our analysis. The app we developed challenged our limitations since our idea was dependent on external IoT devices. Along with the three variables we analyzed, we discovered that it is necessary to include other variables like rainfall, sunlight, soil intensity, vegetation intensity, bees sound frequency, location and topography of the region. Last but not least, we learnt how honeybees are one of the most important organisms of our ecosystem while doing the literature survey.



Fig[11]: <u>Glimpse of our project</u>

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import plotly.express as px
pd.set_option('display.max_rows', 500)
```

Load Dataset

In [2]:
```
beekeeping_df = pd.read_csv('beekeeping_dataset.csv')
```

In [3]:
```
beekeeping_df
```

Out[3]:

	timestamp	temperature	humidity	weight
0	1/3/19 8:0:0 AM	27.3	46.4	3.52
1	1/3/19 8:1:0 AM	27.3	46.2	3.52
2	1/3/19 8:2:0 AM	27.3	46.3	3.52
3	1/3/19 8:3:0 AM	27.5	46.3	3.53
4	1/3/19 8:4:0 AM	27.3	46.3	3.53
...	...	...	...	...
4265	30/4/19 8:5:0 PM	32.8	33.8	5.99
4266	30/4/19 8:6:0 PM	32.8	33.9	5.99
4267	30/4/19 8:7:0 PM	33.5	33.9	5.99
4268	30/4/19 8:8:0 PM	33.5	33.9	5.99
4269	30/4/19 8:9:0 PM	33.5	33.9	5.99

4270 rows × 4 columns

In [4]:
```
beekeeping_df.dtypes
```

Out[4]:
```
timestamp       object
temperature    float64
humidity       float64
weight         float64
dtype: object
```

Frequency Distribution

In [5]:
```
weight_df = beekeeping_df['weight']
humidity_df = beekeeping_df['humidity']
temperature_df = beekeeping_df['temperature']
time_df = beekeeping_df['timestamp']
```

In [6]:
```
weight_df
```

Out[6]:
```
0       3.52
1       3.52
2       3.52
3       3.53
4       3.53
        ...
4265    5.99
4266    5.99
4267    5.99
4268    5.99
4269    5.99
Name: weight, Length: 4270, dtype: float64
```

In [7]:
```
humidity_df
```

Out[7]:
```
0       46.4
1       46.2
2       46.3
3       46.3
4       46.3
        ...
4265    33.8
4266    33.9
4267    33.9
4268    33.9
4269    33.9
Name: humidity, Length: 4270, dtype: float64
```

In [8]:
```
temperature_df
```

Out[8]:
```
0       27.3
1       27.3
2       27.3
3       27.5
4       27.3
        ...
4265    32.8
4266    32.8
4267    33.5
4268    33.5
4269    33.5
Name: temperature, Length: 4270, dtype: float64
```

In [9]:
```
time_df
```

Out[9]:
```
0          1/3/19  8:0:0  AM
1          1/3/19  8:1:0  AM
2          1/3/19  8:2:0  AM
3          1/3/19  8:3:0  AM
4          1/3/19  8:4:0  AM
                ...
4265     30/4/19  8:5:0  PM
4266     30/4/19  8:6:0  PM
4267     30/4/19  8:7:0  PM
4268     30/4/19  8:8:0  PM
4269     30/4/19  8:9:0  PM
Name: timestamp, Length: 4270, dtype: object
```

Counting number of empty rows in humidity, temperature, weight

In [10]:
```
beekeeping_df.isnull().sum()
```

Out[10]:
```
timestamp      0
temperature    4
humidity       5
weight         4
dtype: int64
```

There are total 13 missing values in dataset. There are 4 missing values in temperature, 5 in humidity and 4 in weight feature. ¶

In [11]:
```
clean_data = beekeeping_df.dropna() #removing missing values
```

In [12]:
```
clean_data #printing dataset without missing values
```

Out[12]:

	timestamp	temperature	humidity	weight
0	1/3/19 8:0:0 AM	27.3	46.4	3.52
1	1/3/19 8:1:0 AM	27.3	46.2	3.52
2	1/3/19 8:2:0 AM	27.3	46.3	3.52
3	1/3/19 8:3:0 AM	27.5	46.3	3.53
4	1/3/19 8:4:0 AM	27.3	46.3	3.53
...	...	...	...	...
4265	30/4/19 8:5:0 PM	32.8	33.8	5.99
4266	30/4/19 8:6:0 PM	32.8	33.9	5.99
4267	30/4/19 8:7:0 PM	33.5	33.9	5.99
4268	30/4/19 8:8:0 PM	33.5	33.9	5.99
4269	30/4/19 8:9:0 PM	33.5	33.9	5.99

4258 rows × 4 columns

In [13]:
```
clean_data.isnull().sum() #no missing values left
```

Out[13]:
```
timestamp      0
temperature    0
humidity       0
weight         0
dtype: int64
```

In [14]:
```
clean_data.to_csv('beehive_humidity_temp_weight.csv',index=False)
```

In [15]:
```
beehive_df = pd.read_csv('beehive_humidity_temp_weight.csv')
beehive_df
```

Out[15]:

	timestamp	temperature	humidity	weight
0	1/3/19 8:0:0 AM	27.3	46.4	3.52
1	1/3/19 8:1:0 AM	27.3	46.2	3.52
2	1/3/19 8:2:0 AM	27.3	46.3	3.52
3	1/3/19 8:3:0 AM	27.5	46.3	3.53
4	1/3/19 8:4:0 AM	27.3	46.3	3.53
...	...	...	...	...
4253	30/4/19 8:5:0 PM	32.8	33.8	5.99
4254	30/4/19 8:6:0 PM	32.8	33.9	5.99
4255	30/4/19 8:7:0 PM	33.5	33.9	5.99
4256	30/4/19 8:8:0 PM	33.5	33.9	5.99
4257	30/4/19 8:9:0 PM	33.5	33.9	5.99

4258 rows × 4 columns

Doing inferential statistical analysis, like mean, var, std, percentiles on the give dataset.

In [16]:
```
descr1 = beehive_df["humidity"].describe()
print(descr1)
```

```
count    4258.000000
mean       28.999084
std        13.707917
min        10.100000
25%        18.300000
50%        25.400000
75%        37.375000
max        78.600000
Name: humidity, dtype: float64
```

In [17]:
```
descr2 = beehive_df["temperature"].describe()
print(descr2)
```

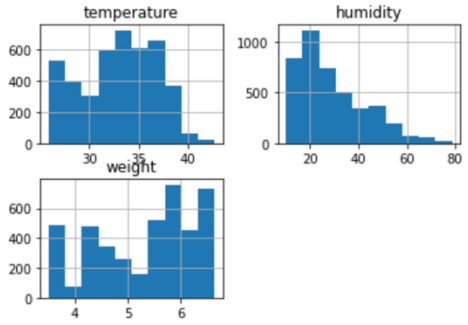
```
count    4258.000000
mean       33.105026
std         3.831551
min        25.900000
25%        29.800000
50%        33.500000
75%        36.200000
max        42.700000
Name: temperature, dtype: float64
```

In [18]:
```
descr3 = beekeeping_df["weight"].describe()
print(descr3)
```

```
count    4266.000000
mean        5.313729
std         0.951502
min         3.510000
25%         4.540000
50%         5.570000
75%         6.140000
max         6.640000
Name: weight, dtype: float64
```

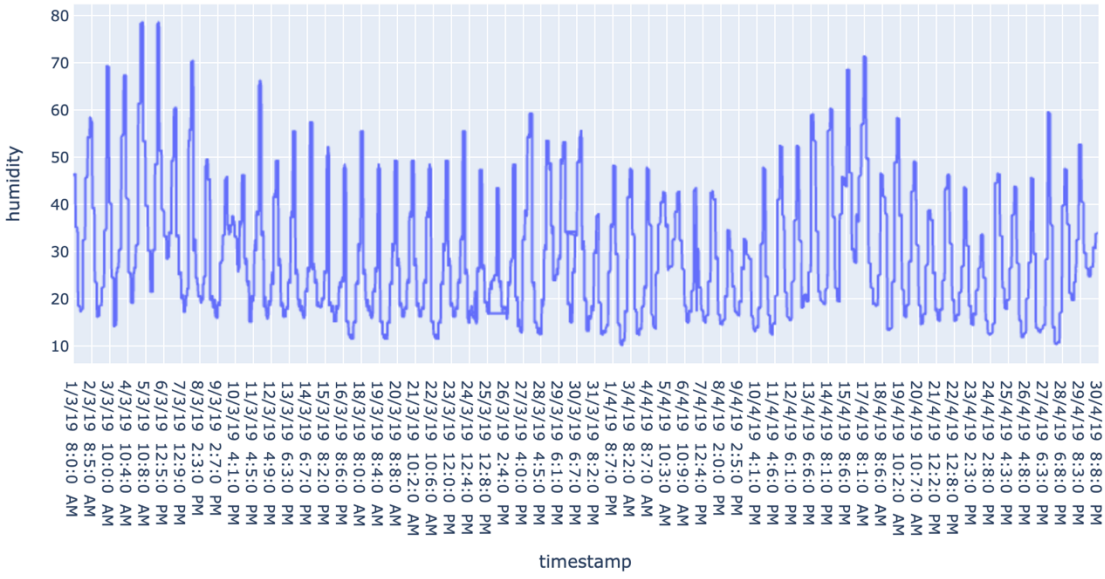
Univariate Graphs

In [19]:
```
beehive_df.hist()
plt.show()
```



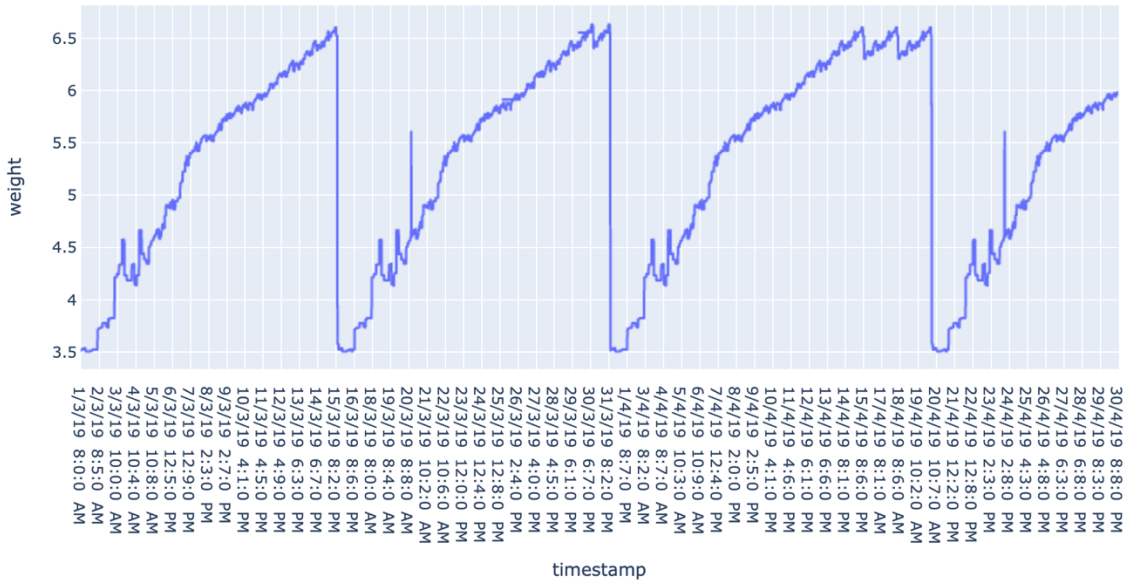
Here we can see humidity is left-skewed or exponential distribution and temperature and weight may be normally distributed.

In [20]:
```
fig = px.line(beehive_df, x='timestamp', y="humidity")
fig.show()
```



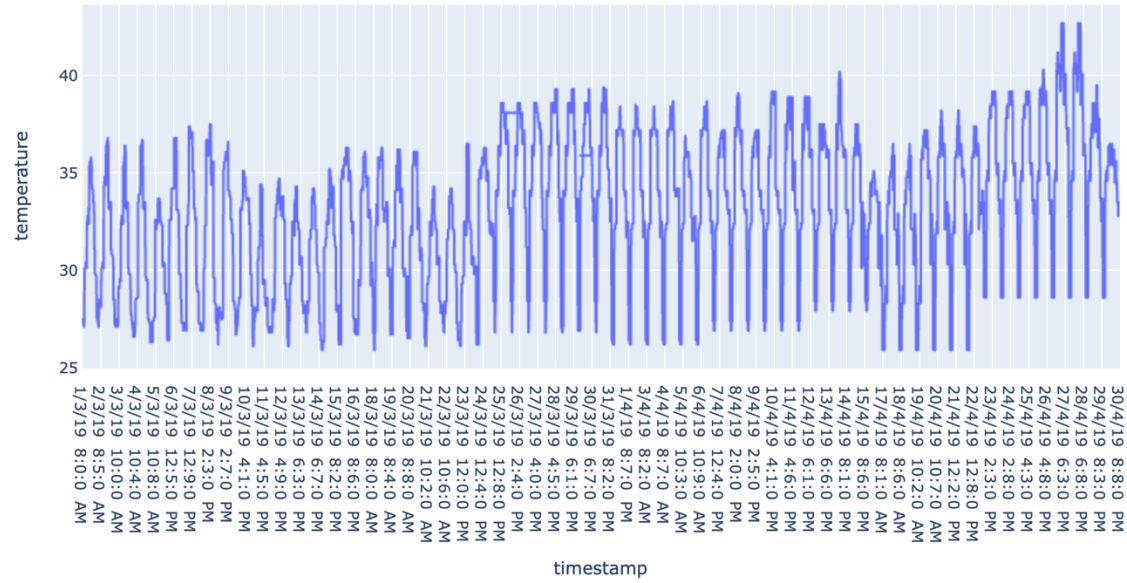
Humidity is high in the morning. It reduces significantly in afternoon at around 2pm. And then increases in the evening

In [21]:
```
fig = px.line(beehive_df, x='timestamp', y="weight")
fig.show()
```



In the month of march the frames in the beehive were emptied twice, once on 16/3/19 and second time on 31/3/19. But in the month of april it was emptied on 20/4/19. So, from this we can infer that yield is higher in month of march than in april.

In [22]:
```
fig = px.line(beehive_df, x='timestamp', y="temperature")
fig.show()
```



Temperature of beehive rises upto a great extent in time period 12-4pm. It goes above threshold value of 32-33 degrees celcius which can affect bees a lot and can result in a bad yield. ¶

In [23]:
```
corr = beehive_df.corr()
corr.style.background_gradient(cmap='coolwarm')
```

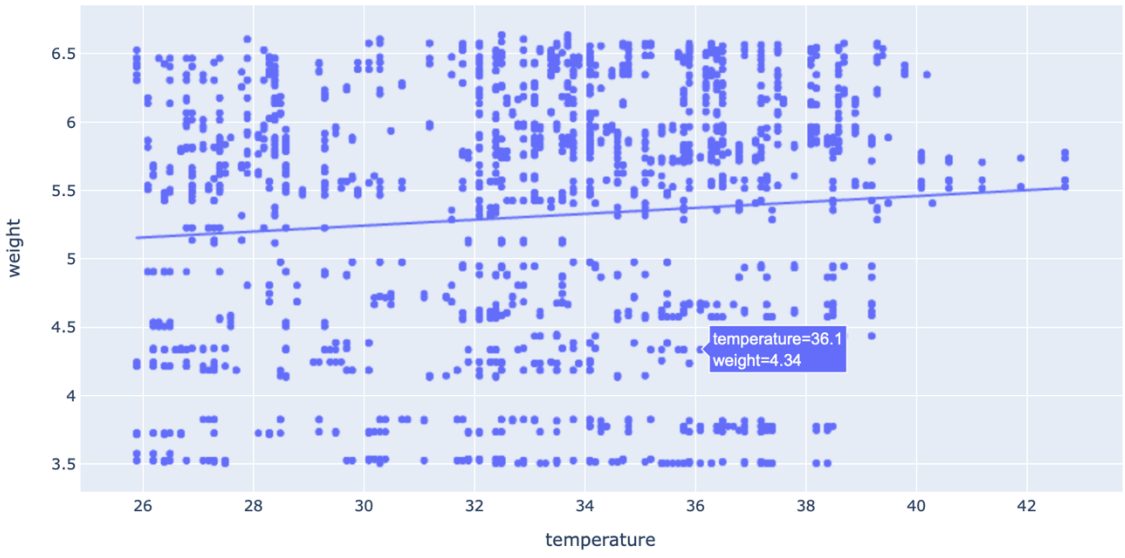
Out[23]:

	temperature	humidity	weight
temperature	1.000000	-0.714077	0.087032
humidity	-0.714077	1.000000	0.057841
weight	0.087032	0.057841	1.000000

Weight increases with increase in temperature. Humidity has less effect on weight

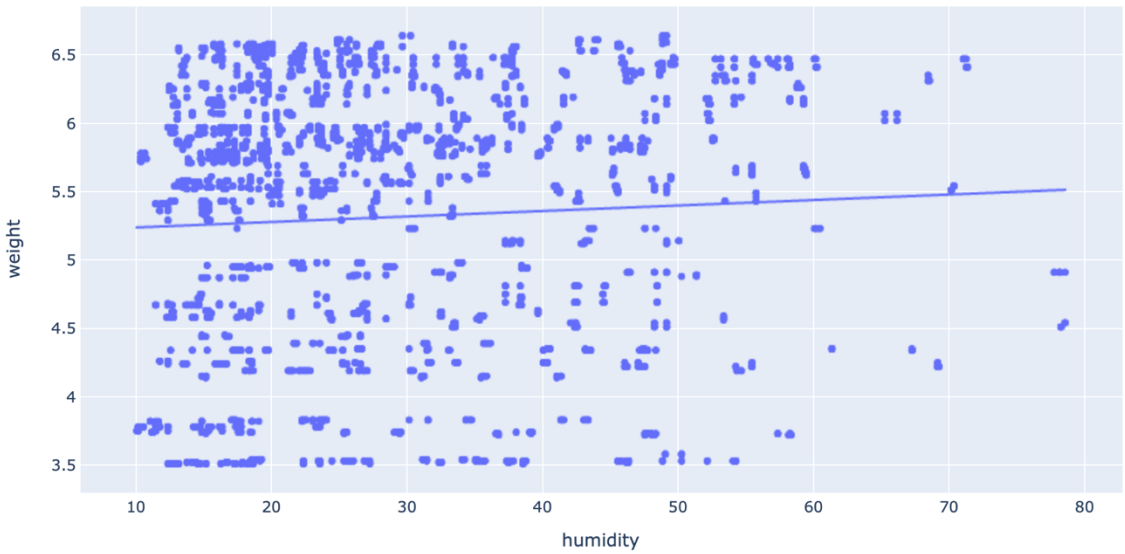
Bivariate Graphs

In [24]:
```
fig = px.scatter(beehive_df, x="temperature", y="weight", trendline="ols")
fig.show()
```



It can be seen that in afternoon when temperature is very high the weight is also high. This is because bees are busy in fanning the beehive.

In [25]:
```
fig = px.scatter(beehive_df, x="humidity", y="weight", trendline="ols")
fig.show()
```



It can be obsereved that when humidity is low the weight is also a bit less but has very weak correlation

7. RECOMMENDATION FOR FUTURE WORK

Since in India beekeepers are not acquainted with smart beekeeping monitoring, this project has a huge scope to undergo development in many areas. The app we developed has static insights and we could improvise to make those insights dynamic. We could further simplify the interface of our app to make beekeepers aware of their actions. We can even include more variables in near future to accurately predict upcoming events. We plan to help beekeepers and make the app useable anywhere in India. Since lots of changes are required, we are thinking to remodel the idea in the backend. Along with apiculture similar data analysis can be upgraded to horticulture and agriculture as well. Since lots of beekeepers collaborate with farmers or manage bee and crop farms individually, we can implement features that will supervise their other activities as well.

8. REFERENCES

1. https://en.wikipedia.org/wiki/Beekeeping

2. https://nbb.gov.in/pdf/Registered%20Beekeepers%202019.pdf

3. http://agritech.tnau.ac.in/farm_enterprises/fe_api_harvestingandprocessing.html

4. http://www.fao.org/home/en/

5. http://www.fao.org/home/en/

6. Research paper – Predictive modelling of honey bee foraging activity using local weather conditions: https://link.springer.com/article/10.1007/s13592-018-0565-3

7. Project repositories - https://github.com/Rahulm0106/Beekeeping-Project.git