# Garmin Run Data

Jenny Kim, Atlanta Liu & Lawrence Shao

# Research Questions

- **Where and when does timestamp mismatch occur in the Garmin dataset?**
- **What patterns can be revealed from using K-Means cluster across all cont. variables? (e.g BPM, Distance, Speed, and Altitude)**


- **How does time spent sleeping affect perceived workout intensity?**
- **For individuals that run the most, do they exhibit any changes in the spread of their heart rate the more they run?**

# Data Cleaning

## Garmin Data

Final dataframe consisted the following columns:

- Subject ID
- Altitude
- Speed
- Datetime
- Distance
- BPM

| | SubjectID | Datetime | Altitude | Distance | Speed | BPM |
|---|---|---|---|---|---|---|
| **0** | 58 | 2018-09-16 00:00:00 | 1117.2 | 3.17 | 1.465 | 78.0 |
| **1** | 58 | 2018-09-16 00:00:02 | 1116.4 | 7.93 | 2.305 | 78.0 |
| **2** | 58 | 2018-09-16 00:00:03 | 1116.2 | 10.39 | 2.454 | 82.0 |
| **3** | 58 | 2018-09-16 00:00:04 | 1116.2 | 12.91 | 2.463 | 86.0 |
| **4** | 58 | 2018-09-16 00:00:05 | 1116.0 | 15.66 | 2.547 | 86.0 |

# Data Cleaning

## Survey Data

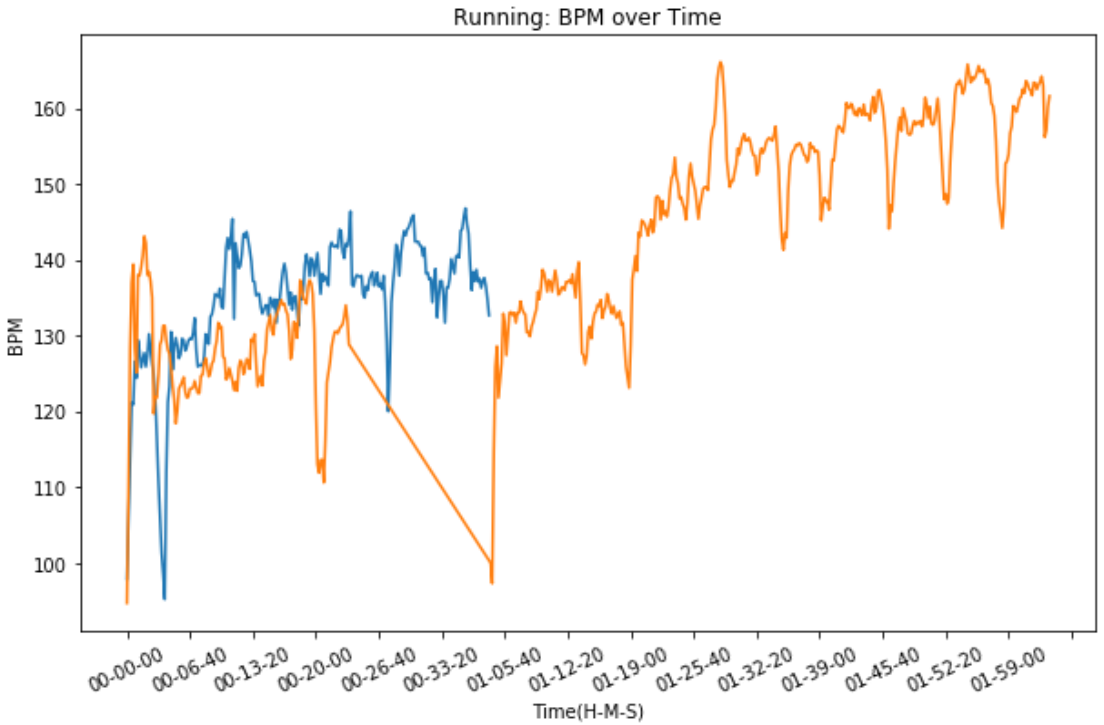
| | SubjectID | EventID | Intensity |
|---|---|---|---|
| **4** | 58 | 2018811 | 3 |
| **9** | 58 | 2018609 | 4 |
| **14** | 58 | 2018526 | 3 |
| **29** | 58 | 2018704 | 4 |
| **35** | 58 | 2018818 | 3 |

| | SubjectID | EventID | Sleep Time |
|---|---|---|---|
| **0** | 58 | 2018811 | 7.5 |
| **5** | 58 | 2018609 | 8.0 |
| **10** | 58 | 2018526 | 7.5 |
| **16** | 58 | 2017722 | 7.5 |
| **25** | 58 | 2018704 | 6.5 |

| | Intensity | Sleep |
|---|---|---|
| **0** | 3 | 7.5 |
| **1** | 3 | 8.0 |
| **2** | 3 | 7.5 |
| **3** | 3 | 7.5 |
| **4** | 3 | 6.5 |

# A Jump in Time

| | SubjectID | Datetime | Altitude | Distance | Speed | BPM | TimeDiff |
|---|---|---|---|---|---|---|---|
| **697511** | 432 | 2017-09-02 00:23:31 | 1276.4 | 3545.84 | 1.521 | 126.0 | 00:00:01 |
| **697512** | 432 | 2017-09-02 00:23:32 | 1276.2 | 3547.45 | 1.521 | 129.0 | 00:00:01 |
| **697513** | 432 | 2017-09-02 00:23:33 | 1276.0 | 3549.15 | 1.521 | 129.0 | 00:00:01 |
| **697514** | 432 | 2017-09-02 00:23:34 | 1275.8 | 3550.93 | 1.521 | 129.0 | 00:00:01 |
| **697515** | 432 | 2017-09-02 00:23:35 | 1275.8 | 3552.63 | 1.530 | 129.0 | 00:00:01 |
| **697516** | 432 | 2017-09-02 01:04:17 | 1264.2 | 3554.61 | 1.642 | 100.0 | 00:40:42 |
| **697517** | 432 | 2017-09-02 01:04:18 | 1263.4 | 3556.99 | 1.885 | 100.0 | 00:00:01 |
| **697518** | 432 | 2017-09-02 01:04:19 | 1263.8 | 3559.52 | 2.258 | 100.0 | 00:00:01 |



# Missing Data

| | Multiple Times In a Run (63 rows) | One Time In a Run (158 rows) |
|---|---|---|
| Median Altitude (m) | 1065 | 1063 |
| Median Distance (m) | 8638 | 3589 |
| Median Speed (m/s) | 0.112 | 1.222 |
| Median BPM | 108 | 100 |
| Median Time Difference (H-M-S) | **00:03:05** | **3:14:50** |

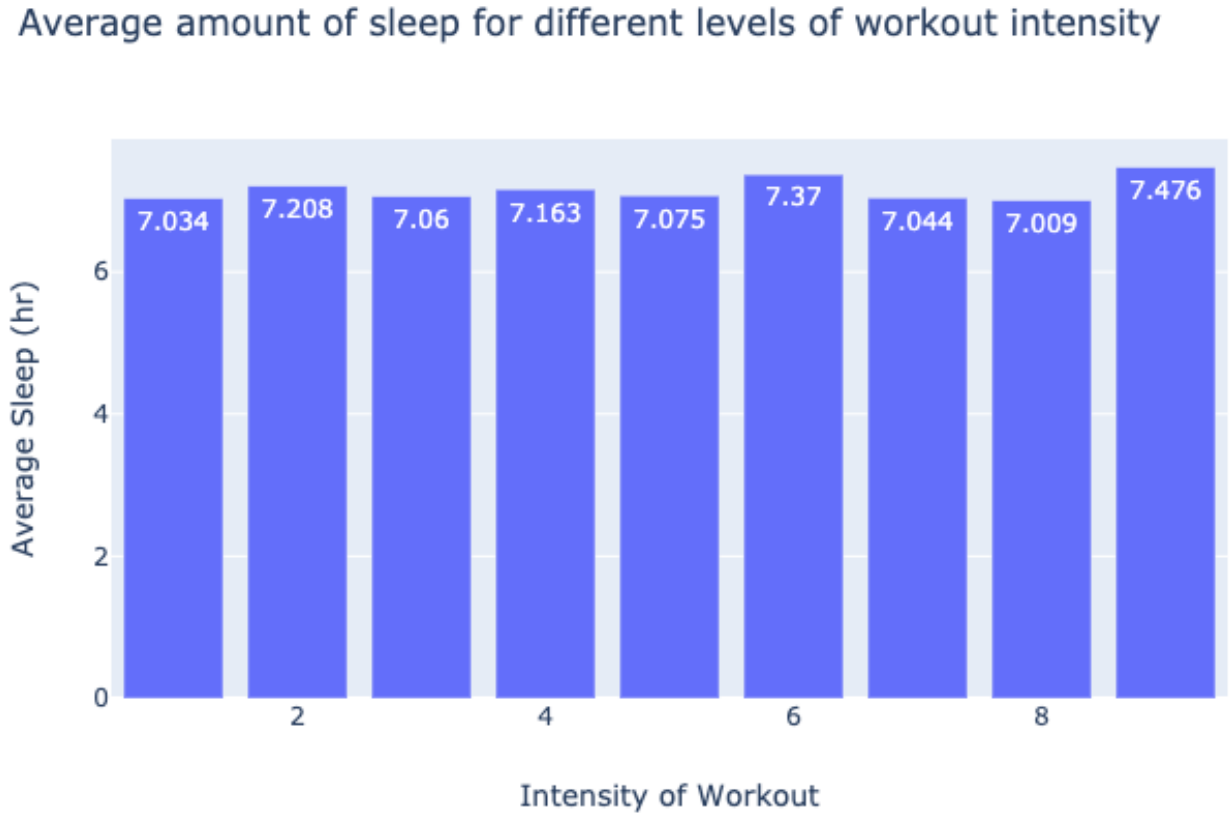
# Results

**K-Means Cluster**

- K-means clustering with 2 clusters was performed
- From the two clusters we can infer that people in higher altitudes tend to run shorter distances when compared to those in lower altitudes


| | Cluster 0 | Cluster 1 |
|---|---|---|
| Mean BPM | 140.27 | 137.57 |
| Mean Altitude | 1094.62 | 149.46 |
| Mean Distance | 7101.39 | 12026.12 |
| Mean Speed | 2.31 | 2.35 |

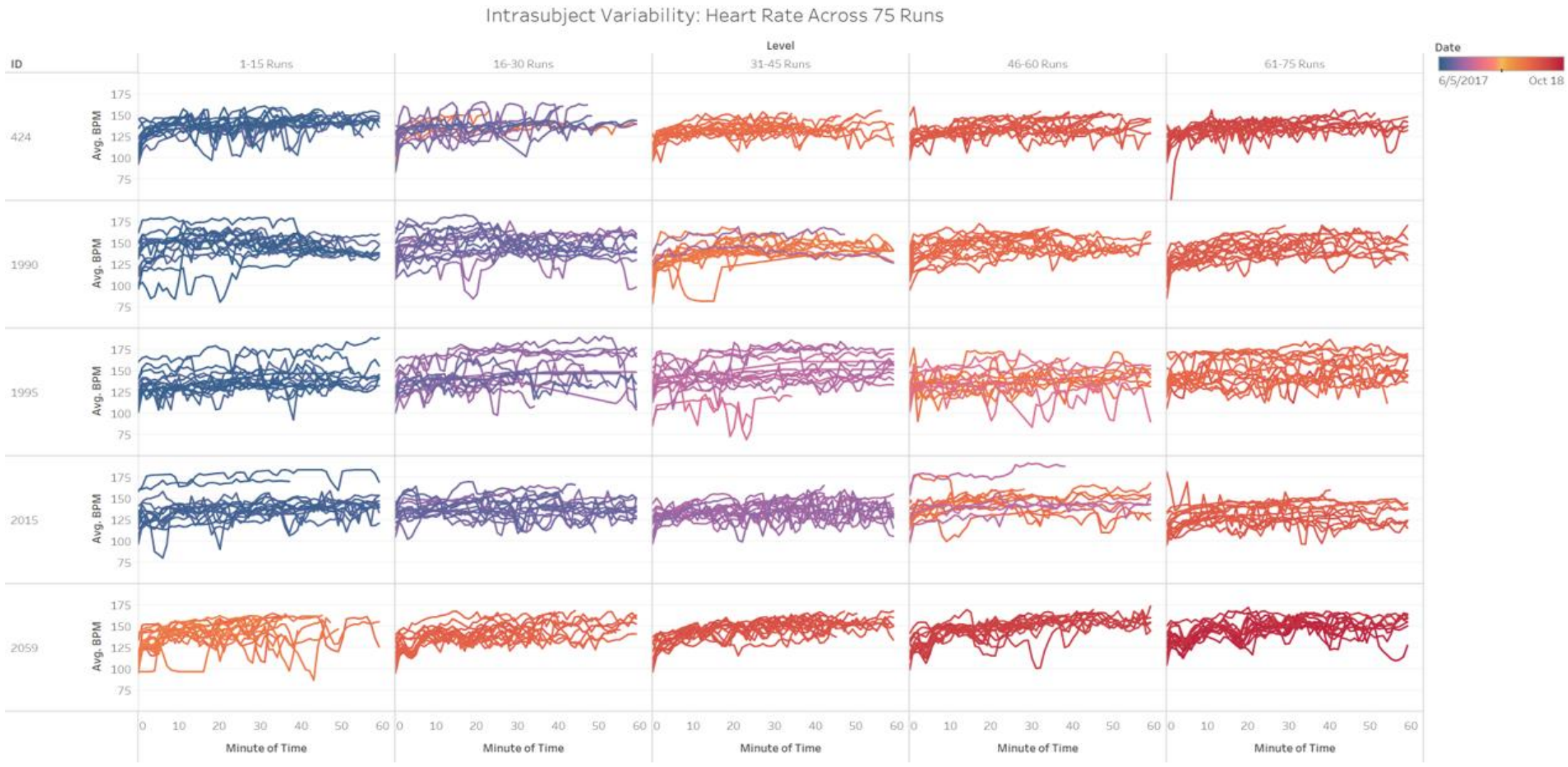
# Results



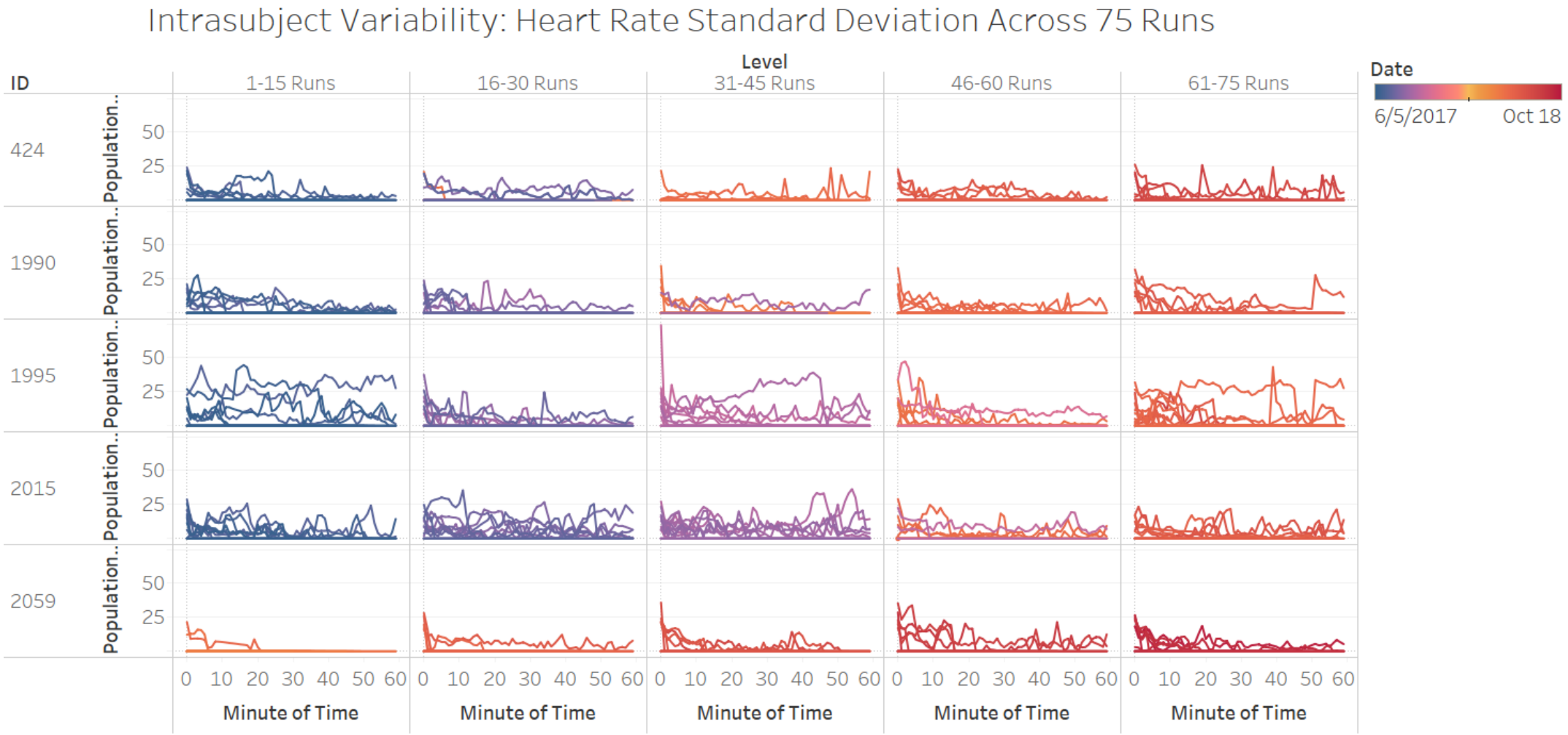

**Associations between sleep and workout intensity:**

- Survey data was used to determine the amount of sleep and workout intensity for every runner
- Average sleep was calculated for each level of workout intensity
- No strong relationship between sleep and workout intensity

# Results


Intrasubject Variability: Heart Rate Across 75 Runs
Level
ID
1-15 Runs
16-30 Runs
31-45 Runs
46-60 Runs
61-75 Runs
Date
6/5/2017
Oct 18
424
1990
1995
2015
2059
Avg. BPM
175
150
125
100
75
0
10
20
30
40
50
60
Minute of Time
The trend of average of BPM for Time Minute broken down by Level vs. ID. Color shows details about Date.

# Results




The trend of population standard deviation of BPM for Time Minute broken down by Level vs. ID. Color shows details about Date.

# Conclusion

**Missing Timestamps**

Shorter timestamp differences if it occurs multiple times on the same day.

**Sleep & Workout Intensity**

No relationships were found between average time spent sleeping and perceived workout intensity

**K-Means Cluster**

Subjects located in higher altitudes will tend to run for shorter distances, whereas individuals in lower altitudes tend to run farther.

**Top Runners**

For individuals that run the most, we did not observe huge changes in the spread of their heart rate as they run

# Further Research & Application

**Heart Rate Recovery & Injury Prediction**

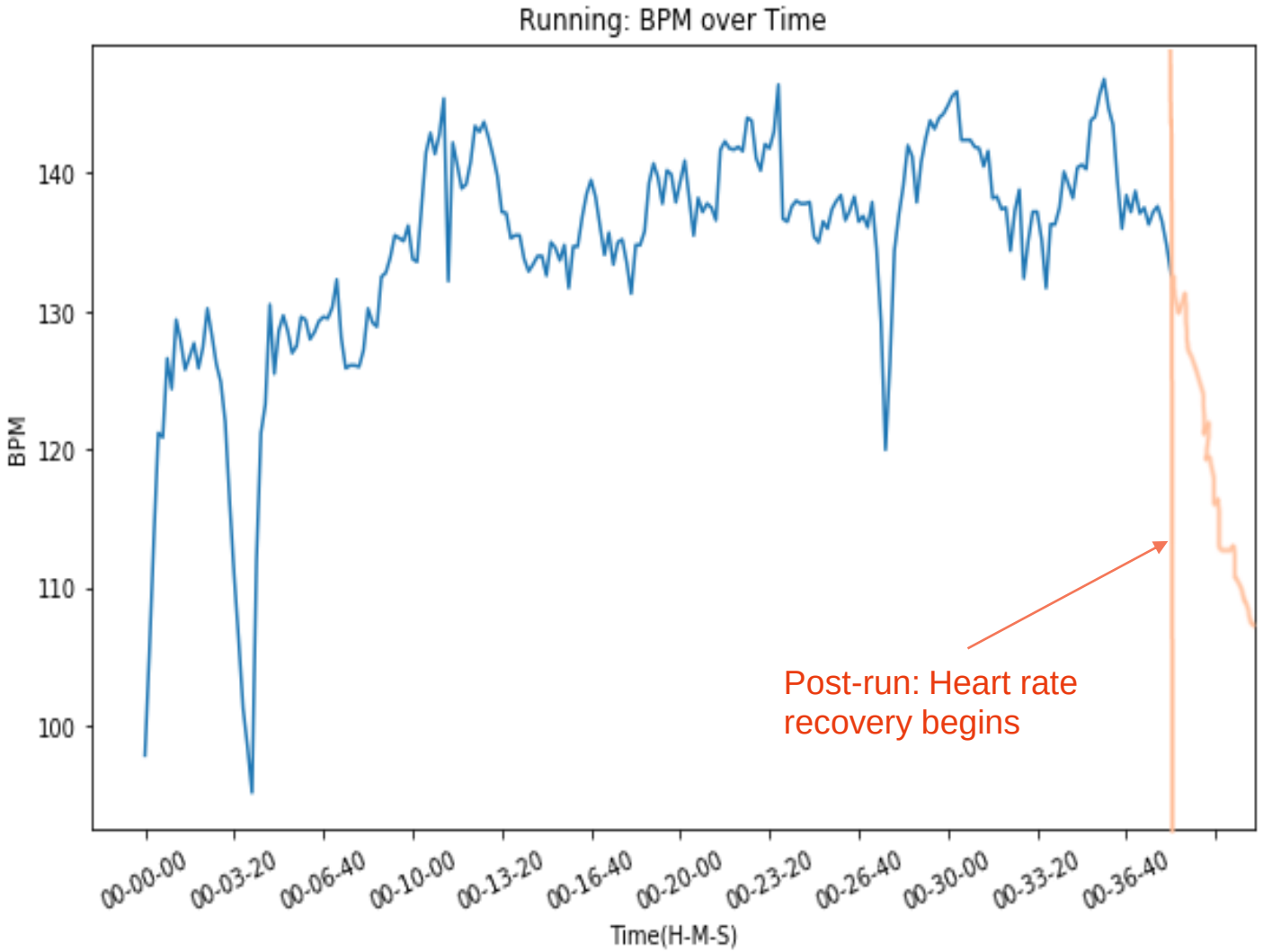
Can data in heart rate recovery be used to improve injury prediction results?

**Implementing a Cooldown Timer**

Instead of stopping all measurements immediately after the end of the run, program it to continuously measure BPM for 1 minute afterwards.

**Understanding Missing Timestamps**

The cooldown timer also serves as a visual indicator to understand when a subject purposefully stops their run.



# THANK YOU