

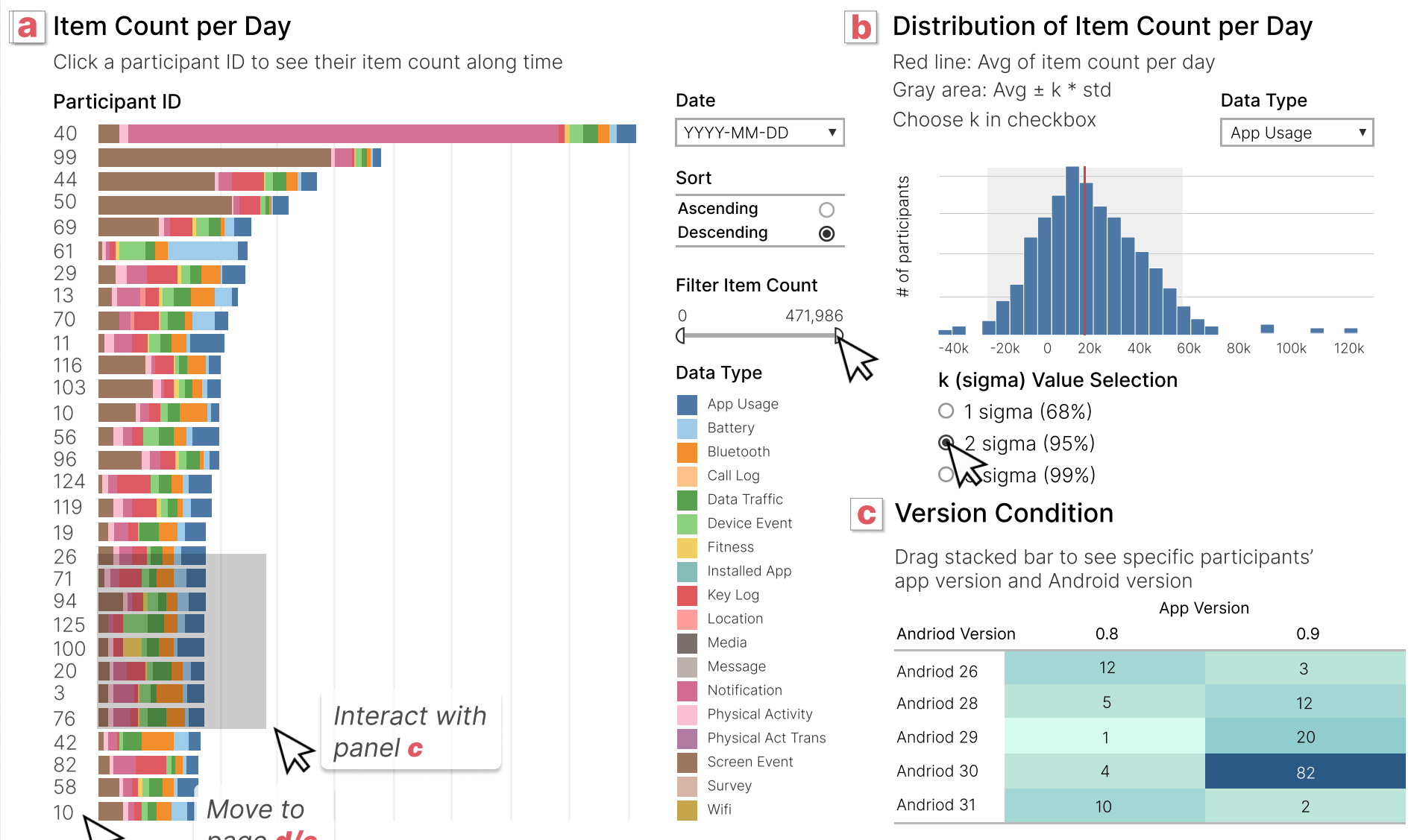
Introduction

- We propose a visualization dashboard to **monitor participants with missing data** in mobile and wearable data collection
- Considering the **event-based sensing characteristic** of mobile and wearable sensor data (e.g., data is logged only when a specific event occurs, such as physical activity or app usage), and calculating **simple quality metric** (item count per day / hour)
- The dashboard provides an visualization to identify participants with **significantly lower metrics compared to others**

System Design

a. Overview panel

- Representing **daily item count per day** for each sensor and person at a glance
- Emphasizing people with **considerably smaller amounts of count** than others



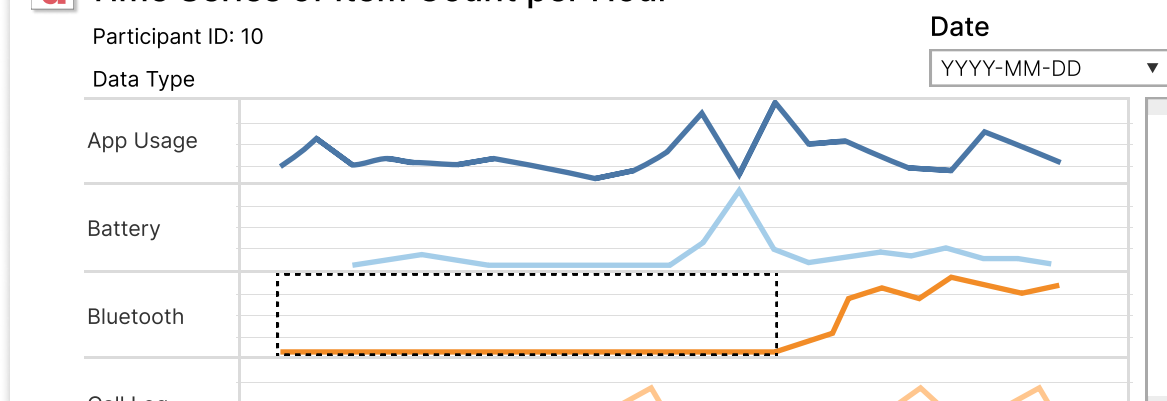
b. Quality control panel

- Using the **concept of control chart** to establish the criteria for identifying problematic item count metric
- Metric value outside **$[\mu - k\sigma, \mu + k\sigma]$** is considered as outliers

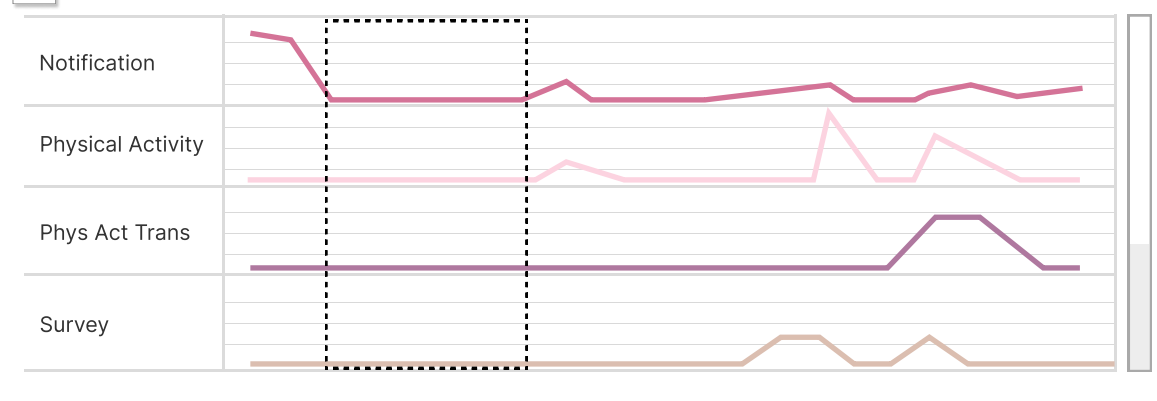
c. Version condition panel

- Observing how many people with small count (who have missing data) use specific **app / Android version**

d Time Series of Item Count per Hour



e



d/e. Time-series exploration page

- Counting sensor data **at an hourly level** to **diagnose the contexts** of missing data collection
 - Missing areas of a **single sensor** (d) due to sensor malfunctions or turning off the sensor
 - Missing areas of **multiple sensor** (e) due to smartphone turning off

Case Study



116 participants



17 kinds of data



1 month



2 researchers

Future Work

- Incorporating **a broader range of data quality issues** such as outliers and abnormal values
- **Patterning common missing data issues** and visualizing them to enable researchers to **diagnose the causes of missing data in more detail**