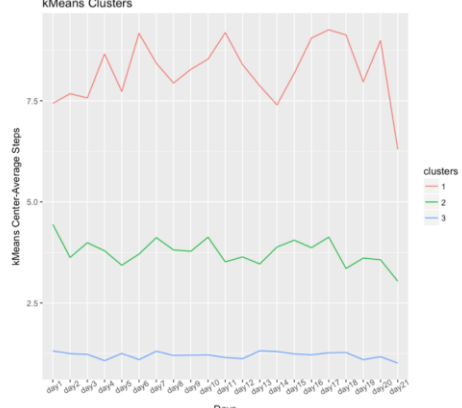
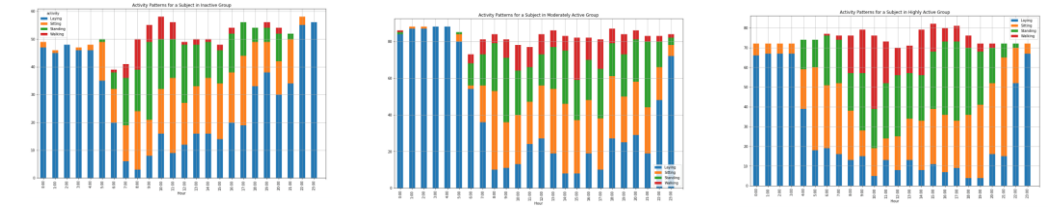
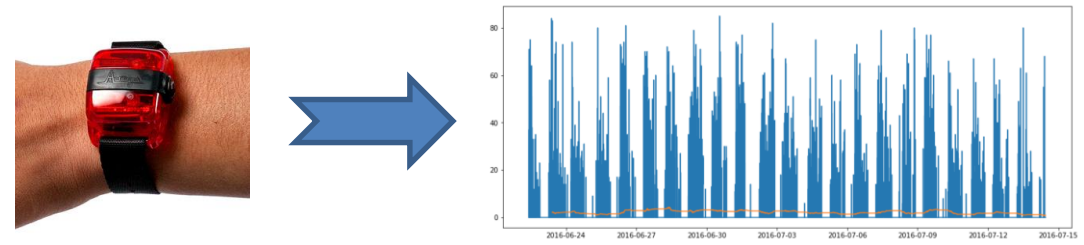
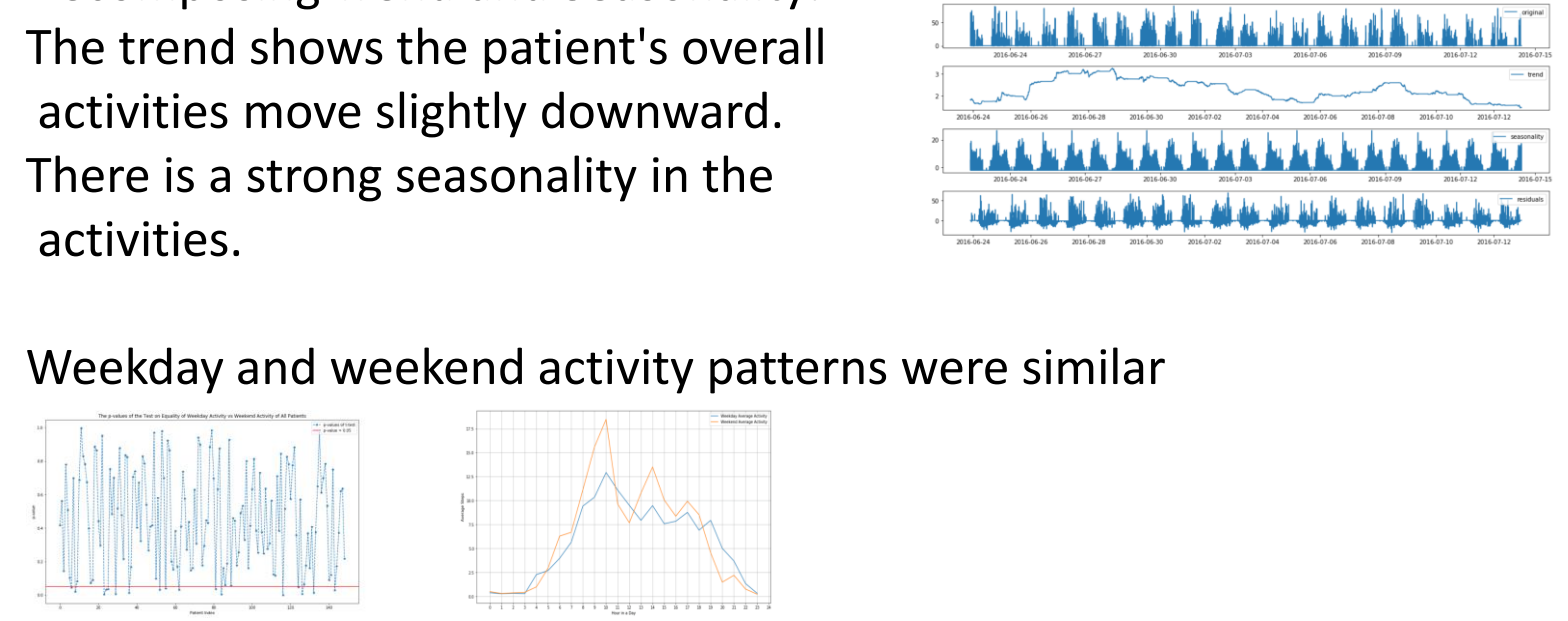


7a.007.DU - Novel Methods for Data Integration and Information Extraction over Complex Heterogeneous Data

Yuan An¹ (PI), Siling Chen¹ (RA), Matthew Marschall¹(RA), Nicholas Locantore²(IAB), Matthew Allinder²(IAB)
 College of Computing and Informatics, Drexel University¹, GSK²

Project Start: 07/01/2018	End Date: 06/30/2019	Project Budget: \$30,000	Spent:	<ul style="list-style-type: none"> Clustering the patients into three types of activities: <ul style="list-style-type: none"> Highly Intensive (class 1) Moderately Intensive (class 2) Low Intensive (class 3) 
Project Summary: <ul style="list-style-type: none"> Developing innovative big data approach for extracting, integrating, visualizing, and analyzing big data sets and for helping researchers gain knowledge from the large amount of data related to business, health care, and scientific research activities. In particular, investigating and developing data analytic methods for discovering and Visualizing Physical Activity Patterns for COPD Patients in Real-time Digital Monitoring Data. 				<ul style="list-style-type: none"> The automatically labeled activities: Blue: Laying; Orange: Sitting; Green: Standing; Red: Walking for an example patient in low, medium, and high intensity group. 
Details of Progress/Achievements: <ul style="list-style-type: none"> Use of portable digital monitoring devices to continuously collect physical activity and patient reported outcomes are becoming more common. There is a need to effectively assess activity patterns and manage disease risks for COPD patients. We analyze the data about 183 individuals from a single center of the COPDGene cohort in a three-week observational study. Each participant used a smartphone to complete a daily symptom diary, wore a wrist-worn accelerometer to continuously record physical activity, and completed an activity questionnaire. The problems are: <ul style="list-style-type: none"> Are there any salient patterns in physical activity useful for disease assessment and monitoring? Discover and compare the physical activity patterns for different cohorts of patients 				<ul style="list-style-type: none"> Decomposing Trend and Seasonality: <ul style="list-style-type: none"> The trend shows the patient's overall activities move slightly downward. There is a strong seasonality in the activities. Weekday and weekend activity patterns were similar 

PROJECT DELIVERABLES

Deliverable	Achievements	Remaining To Do
Analyzing and visualizing patient activity data for discovering patterns for individual and cohort patients	Analyzed all 183 patients' physical activity data and created visualization models for understanding patients' behaviors	Detecting abnormal behaviors during different time periods for disease exacerbation analysis
Linking the physical activity data with clinical data and applying advanced data mining methods for discovering any correlations between physical activities and clinical progressions	Applied associate rule mining methods on combined physical activity and clinical data by creating appropriate items through analyzing the key features of the data	Detecting shapelets associated with key clinically defined events such as inhaler uses and disease progression trends
Publications in relevant scientific research conferences or journals.	A poster has been accepted and accepted by the 2018 MidAtlantic Bioinformatics Conference in Philadelphia	Submit abstracts and full research papers to relevant conferences and journals