В ходе первой сессии данные были обработаны и от временных рядов для каждого параметра был осуществлен переход к некоторым фичам, которые не зависят от длительности расчета показателей пациента.

В ходе второй сессии:

- Из данных были убраны дубликаты и пустые столбцы
- Был убран параметр MeanMechVent, т.к. он состоял из одних единиц и пропусков
- Параметры Gender, Height, MeanWeight, ICUType были приведены к правильным значениям (в данных присутствовали отрицательный вес, гендер=0.5 и прочее)
- Были удалены столбцы, в которых больше половины пропусков
- Остальные пропуски были заменены средним значением столбца
- Были удалены выбросы используя IsolationForest
- Были построены гистограммы параметров, боксплоты и коррелограмма, в результате анализа последней было получено, что заметна высокая корреляция
  признаков MeanNIMAP с MeanNiDiasABP и MeanNIMAP с MeanNiSysABP

В ходе третьей сессии:

- Данные были нормализованы
- Из датасета были получены два отдельных датасета (для задачи прогноза продолжительности жизни и для задачи предсказания выживаемости)
- В результате отбора признаков (отбор проводился используя обучение моделей с L1-регуляризатором, который зануляет коэффициенты перед малозначимыми признаками) в каждом датасете остались по 20 признаков из 34
- Поиск зависимостей в данных проводился используя коррелограмму Пирсона и Кендалла. Также зависимости были оценены визуально (построены парные графики всех зависимостей)
- В данных для классификации присутствовал дисбаланс классов (6328 наблюдений класса 0 против 872 наблюдений класса 1). Данные для тренировки и обучения делились в отношении 80/20. При обучении моделей использовалась кросс-валидация
- Для поиска лучшей регрессионной модели были рассмотрены линейная регрессия, Lasso-регрессия, Ridge-регрессия, ElasticNet и RandomForest. В результате лучшей регрессионной моделью оказалась Lasso-регрессия
- Для поиска лучшей классификационной модели были рассмотрены логистическая регрессия, случайный лес и SVM. В результате лучшей классификационной моделью оказался SVM (ROC-AUC = 0.74)

В ходе четвертой сессии было написано консольное приложение для:

- Прогноза продолжительности жизни пациентов
- Предсказания выживаемости пациента
- Вывода индикатора степени тяжести больного

Приложение делает предсказания по данным из файла с информацией о пациентах и позволяет получить результат тремя разными способами:

1. Вывести его на экран
2. Получить в формате json из скрипта
3. Сохранить таблицу с результатом в файл

Приложение подгружает предобученные модели, обрабатывает данные, делает по обработанным данным предсказание и создает датафрейм с предсказаниями. Далее делает с датафреймом то, что указал пользователь (вывод, возвращение, сохранение)