Набор данных представляет из себя зарегистрированные при помощи датчиков реакции на вопросы о факторах риска на проверках при трудоустройстве. Они размеченны специалистами по трем уровням стресса (слабый, средний, сильный). Каждая проверка представляет из себя несколько групп вопросов, каждый из которых задается не более 4 раз. Для каждого человека была следующая структура тестирования:

Группа1 Повторение1: Вопрос1, Вопрос2, ..., Вопрос12

Группа1 Повторение2: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения1 в случайном порядке)

Группа1 Повторение3: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения1 в случайном порядке)

Группа2 Повторение1: Вопрос1, Вопрос2, ..., Вопрос12

Группа2 Повторение2: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения2 в случайном порядке)

Группа2 Повторение3: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения2 в случайном порядке)

и так далее ...

Количество групп от 1 до 7, в группах может быть от 1 до 14 вопросов.

Соответственно, в наборе данных присутствуют:

Filename - уникальный id испытуемого.

Test\_index - номер группы, нумерация с 0

Presentation - номер повторения, нумерация с 1

Question - номер вопроса, нумерация с 0

Data - равномерно зарегистрированные полиграфом 12 секунд сигнала с частотой 20 Гц (240 значений) (метка датчика регистрации сигнала - фотоплетизмограмма)

Data\_2 - равномерно зарегистрированные полиграфом 12 секунд сигнала с частотой 20 Гц (240 значений) (метка датчика регистрации сигнала - пьезоплетизмограмма)

Class\_label Data\_type - метка класса (0 - слабый стресс, 1 - средний стресс, 2 - сильный стресс)

Важное примечание: согласно методике полиграфолог при выставлении баллов оценивает реакции на вопросы при помощи попарного сравнения в рамках каждого повторения отдельно - не более 14 вопросов. Такая процедура выполняется независимо по каждой группе вопросов по каждому повторению.