

There are 14 columns in the dataset, where the `patient_id` column is a unique and random identifier. The remaining 13 features are described in the section below.

There are 14 columns in the dataset, where the `patient_id` column is a unique and random identifier. The remaining 13 features are described in the section below.

- `slope_of_peak_exercise_st_segment` (type: int): the slope of the peak exercise [ST segment](#), an electrocardiography read out indicating quality of blood flow to the heart
- `thal` (type: categorical): results of [thallium stress test](#) measuring blood flow to the heart, with possible values `normal`, `fixed_defect`, `reversible_defect`
- `resting_blood_pressure` (type: int): resting blood pressure
- `chest_pain_type` (type: int): chest pain type (4 values)
- `num_major_vessels` (type: int): number of major vessels (0-3) colored by flourosopy
- `fasting_blood_sugar_gt_120_mg_per_dl` (type: binary): fasting blood sugar > 120 mg/dl
- `resting_ekg_results` (type: int): resting electrocardiographic results (values 0,1,2)
- `serum_cholesterol_mg_per_dl` (type: int): serum cholestoral in mg/dl
- `oldpeak_eq_st_depression` (type: float): oldpeak = [ST depression](#) induced by exercise relative to rest, a measure of abnormality in electrocardiograms
- `sex` (type: binary): `0`: female, `1`: male
- `age` (type: int): age in years
- `max_heart_rate_achieved` (type: int): maximum heart rate achieved (beats per minute)
- `exercise_induced_angina` (type: binary): exercise-induced chest pain (`0`: False, `1`: True)

• Submission format

-
- The format for the submission file is two columns with the `patient_id` and `heart_disease_present`. This competition uses log loss as its evaluation metric, so the `heart_disease_present` values you should submit are the **probabilities that a patient has heart disease** (not the bin