h1_temp_max	0.236957
h1_temp_min	0.236957
hospital_admit_source	0.233435
hematocrit_apache	0.216741
bun_apache	0.210025
creatinine_apache	0.205565
sodium_apache	0.202807

Look for columns that are mostly empty

> 90% : bilirubin, lactate, albumin
> 80 - 90% : h1_pao2fio2ratio, arterial_ph, hco3, arterial pco2, wbc, arterial_po2, calcium, platelets,bun, creatinine, diasbp_invasive, sysbp_invasive, mbp, invasive, hematocrit
> 70 - 80% : hemaglobin, sodium, potassium, ph_apache, pao2_apache, paco2_for_ph_apache, fio2_apache, paco2_apache, diasbp_invasive, sysbp_invasive, mbp_invasive, d1_pao2fio2ratio
> ** Note ** all of the apache scores have the same missing values
> 60 - 70% : arterial_ph, arterial_pco2, arterial_po2, bilirubin_apache, inr both h1 and d1
> 50 - 60% : albumin_apache, bilirubin, glucose, albumin, urineoutput_apache
> 30 - 50% : no columns
> 20 - 30% : wbc_apache, temp, hosptial_admit_source, hematocrit_apache, bun_apache, creatinine_apache, sodium_apache
> 10 - 20% : all d1 vitals: hco3, platelets, wbc, calcium, hemaglobin, hematocrit, glucose apache, bun, sodium, creatininte, potassium

Empty columns can be grouped

- Top missing data is from h1_vitals. Here 'h1' means it was taken during their first hour in the stay. 'd1' is during the first 24 hours.
  - It could be intersting to look if the patients have both h1 and d1, which may mean two measurements. Check if they are the same.
  - It could also be that some hospitals take vitals within the hour or that having an h1 score at all means a condition is more serious