

Data dictionary

1. **'Patient ID'**-unique ID identifying an anonymized patient
2. **'Patient age quantile'**- bins representing the patients age, as interpreted on kaggle by others
 - 1: 0-5
 - 2: 6-10
 - 3: 11-15
 - 4: 16-20
 - 5: 21-25
 - 6: 26-30
 - 7: 31-35
 - 8: 36-40
 - 9: 41-45
 - 10: 46-50
 - 11: 51- 55
 - 12: 56-60
 - 13: 61-65
 - 14: 66-70
 - 15: 71-75
 - 16: 76-80
 - 17: 81-85
 - 18: 86-90
 - 19: 91-95
3. **'SARS-Cov-2 exam result'**-Positive or negative PCR result, 558 positive results, categorical variable
4. **'Patient admitted to regular ward (1=yes, 0=no)'** self-explanatory
5. **'Patient admitted to semi-intensive unit (1=yes, 0=no)'**-semi-intensive or intermediate care unit, is usually the place to move improving ICU patients or deteriorating regular ward patients. The semi-intensive unit can be for a patient who is bad but not deteriorating rapidly.
6. **'Patient admitted to intensive care unit (1=yes, 0=no)'**- self explanatory, for the most critical, on the edge of death patients, ICU beds typically cost between \$25,000 and \$30,000. The cost of an ICU bed per night is \$1,107, according to a recent study of two Washington hospitals.
7. **'Hematocrit'** -volume percentage of red blood cells in blood, measured as part of a blood test. The measurement depends on the number and size of red blood cells. It is normally 40.7%-50.3% for males and 36.1%-44.3% for females.
8. **'Hemoglobin'** -abbreviated Hb or Hgb, is the iron-containing oxygen-transport metalloprotein in the red blood cells of almost all vertebrates as well as the tissues of some invertebrates. Hemoglobin in blood carries oxygen from the lungs or gills to the rest of the body.

9. **'Platelets'**- or thrombocytes, are small, colorless cell fragments in our blood that form clots and stop or prevent bleeding. Platelets are made in our bone marrow, the sponge-like tissue inside our bones. Bone marrow contains stem cells that develop into red blood cells, white blood cells, and platelets.
10. **'Mean platelet volume'**-Platelets are small blood cells that are essential for blood clotting, the process that helps you stop bleeding after an injury. An MPV blood test measures the average size of your platelets. The test can help diagnose bleeding disorders and diseases of the bone marrow. A high MPV means that your platelets are larger than average. This is sometimes a sign that you're producing too many platelets. Platelets are produced in the bone marrow and released into the bloodstream. Larger platelets are usually young and more recently released from the bone marrow.
11. **'Red blood Cells'**- haematids, erythroid cells or erythrocytes, are the most common type of blood cell and the vertebrate's principal means of delivering oxygen to the body tissues—via blood flow through the circulatory system.
12. **'Lymphocytes'**- a type of white blood cell in the immune system of jawed vertebrates. Lymphocytes include natural killer cells, T cells, and B cells. They are the main type of cell found in lymph, which prompted the name "lymphocyte".
13. **'Mean corpuscular hemoglobin concentration\xa0(MCHC)'**- It's a measure of the average concentration of hemoglobin inside a single red blood cell. MCHC is commonly ordered as part of a complete blood count (CBC) panel.
14. **'Leukocytes'**- A type of blood cell that is made in the bone marrow and found in the blood and lymph tissue. Leukocytes are part of the body's immune system. They help the body fight infection and other diseases.
15. **'Basophils'**-a type of white blood cell. Like most types of white blood cells, basophils are responsible for fighting fungal or bacterial infections and viruses. They are a granulocyte cell, which means that they release granules of enzymes to fight against harmful bacteria and germs.
16. **'Mean corpuscular hemoglobin (MCH)'**- the average amount in each of your red blood cells of a protein called hemoglobin, which carries oxygen around your body.
17. **'Eosinophils'**-sometimes called eosinophiles or, less commonly, acidophils, are a variety of white blood cells and one of the immune system components responsible for combating multicellular parasites and certain infections in vertebrates.
18. **'Mean corpuscular volume (MCV)'**- or mean cell volume, is a measure of the average volume of a red blood corpuscle. The measure is obtained by multiplying a volume of blood by the proportion of blood that is cellular, and dividing that product by the number of erythrocytes in that volume.

19. **'Monocytes'**-a type of leukocyte, or white blood cell. They are the largest type of leukocyte and can differentiate into macrophages and conventional dendritic cells. As a part of the vertebrate innate immune system monocytes also influence the process of adaptive immunity.
20. **'Red blood cell distribution width (RDW)'**-a measurement of the range in the volume and size of your red blood cells (erythrocytes). Red blood cells move oxygen from your lungs to every cell in your body. Your cells need oxygen to grow, reproduce, and stay healthy. If your red blood cells are larger than normal, it could indicate a medical problem
21. **'Serum Glucose'**-the simplest and most direct single test available to test for diabetes. The test measures the amount of glucose in the fluid portion of the blood. It is called a "serum" test due to this fluid portion of the blood.
22. **'Respiratory Syncytial Virus'**-Respiratory syncytial (sin-SISH-uhl) virus, or RSV, is a common respiratory virus that usually causes mild, cold-like symptoms. Most people recover in a week or two, but RSV can be serious, especially for infants and older adults. RSV is the most common cause of bronchiolitis (inflammation of the small airways in the lung) and pneumonia (infection of the lungs) in children younger than 1 year of age in the United States.
23. **'Influenza A'**-Type A influenza is generally considered worse than type B influenza. Causes pandemics, found in many types of birds, hence bird flu.
24. **'Influenza B'**- Unlike type A flu viruses, type B flu is found only in humans. Type B flu may cause a less severe reaction than type A flu virus, but occasionally, type B flu can still be extremely harmful. Influenza type B viruses are not classified by subtype and do not cause pandemics.
- 25.
26. **'Parainfluenza 1'**-Human parainfluenza viruses (HPIVs) commonly cause upper and lower respiratory illnesses in infants, young children, older adults, and people with weakened immune systems, but anyone can get infected. After you get infected, it takes about 2 to 7 days before you develop symptoms. Symptoms of upper respiratory illness may include fever, runny nose, cough. You can have multiple HPIV illnesses in your lifetime.
27. **'CoronavirusNL63'**- Respiratory tract infection is a leading cause of morbidity and mortality worldwide, especially among young children. Human coronaviruses (HCoVs) have only recently been shown to cause both lower and upper respiratory tract infections.
28. **'Rhinovirus/Enterovirus'**- Rhinoviruses (RVs) and respiratory enteroviruses (EVs) are leading causes of upper respiratory tract infections and among the most frequent infectious agents in humans worldwide. Both are classified in the Enterovirus genus within the

Picornaviridae family and they have been assigned to seven distinct species, RV-A, B, C and EV-A, B, C, D. As viral infections of public health significance, they represent an important financial burden on health systems worldwide. However, the lack of efficient antiviral treatment or vaccines against these highly prevalent pathogens prevents an effective management of RV-related diseases.

29. **'Mycoplasma pneumoniae'**- is an 'atypical' bacterium that causes lung infection. It is a common cause of community-acquired pneumonia
30. **'Coronavirus HKU1'**-Human coronavirus HKU1 is a species of coronavirus in humans. It causes an upper respiratory disease with symptoms of the common cold, but can advance to pneumonia and bronchiolitis. It was first discovered in January 2004 from one man in Hong Kong.
31. **'Parainfluenza 3'**-
32. Parainfluenza virus type 3 is one of a group of common viruses known as human parainfluenza viruses (HPIV) that cause a variety of respiratory illnesses. Symptoms usually develop between 2 and 7 days from the time of exposure and typically resolve in 7-10 days. Symptoms may include fever, runny nose, and cough.
33. **'Chlamydia pneumoniae'**- Chlamydia pneumoniae is a type of bacteria that can cause respiratory tract infections, such as pneumonia. C. pneumoniae is one cause of community-acquired pneumonia or lung infections developed outside of a hospital. However, not everyone exposed to C. pneumoniae will develop pneumonia.
34. **'Adenovirus'**-common viruses that cause a range of illness. They can cause cold-like symptoms, fever, sore throat, bronchitis, pneumonia, diarrhea, and pink eye (conjunctivitis). You can get an adenovirus infection at any age. People with weakened immune systems or existing respiratory or cardiac disease are more likely than others to get very sick from an adenovirus infection
35. **'Parainfluenza 4'**-The parainfluenza viruses are paramyxoviruses and classified as types 1, 2, 3, and 4. They share antigenic cross-reactivity but tend to cause diseases of different severity. Type 4 has antigenic cross-reactivity with the mumps virus and is an uncommon cause of respiratory disease that requires medical attention.
36. **'Coronavirus229E'**-a species of coronavirus which infects humans and bats. It is an enveloped, positive-sense, single-stranded RNA virus which enters its host cell by binding to the APN receptor. Along with Human coronavirus OC43, it is one of the viruses responsible for the common cold.
37. **'CoronavirusOC43'**-infects humans and cattle.The infecting coronavirus is an enveloped, positive-sense, single-stranded RNA virus which enters its host cell by binding to the N-acetyl-9-O-acetylneuraminic acid receptor. OC43 is one of seven

known coronaviruses to infect humans. It is one of the viruses responsible for the common cold.

38. **'Inf A H1N1 2009'**,The virus has caused disproportionate disease among young people with early reports of virulence similar to that of seasonal influenza.
39. **'Bordetella pertussis'**, a Gram-negative, aerobic, pathogenic, encapsulated coccobacillus of the genus Bordetella, and the causative agent of pertussis or whooping cough. Like B. bronchiseptica, B. pertussis is motile and expresses a flagellum-like structure.[1] Its virulence factors include pertussis toxin, adenylate cyclase toxin, filamentous haemagglutinin, pertactin, fimbria, and tracheal cytotoxin.
40. **'Metapneumovirus'**, a common respiratory virus that causes an upper respiratory infection (like a cold). It is a seasonal disease that usually occurs in the winter and early spring, similar to RSV and the flu.
41. **'Parainfluenza 2'**,Parainfluenza viruses belong to the Paramyxoviridae viral family and they have a medium-sized virion particle, are enveloped, nonsegmented, and their genome is composed of a negative-strand RNA. This serotype of parainfluenza belongs to the genus Rubulavirus.
42. **'Neutrophils'**,a type of white blood cell. They make up the biggest number of all kinds of white blood cells. They kill and digest bacteria and fungi to help your body fight infections and heal wounds
43. **'Urea'**,is the chief nitrogenous end product of the metabolic breakdown of proteins in all mammals and some fishes. The material occurs not only in the urine of all mammals but also in their blood, bile, milk, and perspiration.
44. **'Proteina C reactiva mg/dL'**,A CRP test measures the amount of CRP in the blood to detect inflammation due to acute conditions or to monitor the severity of disease in chronic conditions.(0.8 to 100 mg/dL)
45. **'Creatinine'**, a waste product produced by muscles from the breakdown of a compound called creatine.
46. **'Potassium'**, blood potassium level is 3.6 to 5.2 millimoles per liter (mmol/L). A very low potassium level (less than 2.5 mmol/L) can be life-threatening and requires urgent medical attention.
47. **'Sodium'**, A normal blood sodium level is between 135 and 145 milliequivalents per liter (mEq/L). Hyponatremia occurs when the sodium in your blood falls below 135 mEq/L
48. **'Influenza B, rapid test'**, is the only species in the genus Betainfluenzavirus in the virus family Orthomyxoviridae. Influenza B virus is known only to infect humans and seals
49. **'Influenza A, rapid test'**, A virus causes influenza in birds and some mammals, and is the only species of the genus

Alphainfluenzavirus of the virus family Orthomyxoviridae. Strains of all subtypes of influenza A virus have been isolated from wild birds

50. **'Alanine transaminase'**, a transaminase enzyme. It is also called alanine aminotransferase and was formerly called serum glutamate-pyruvate transaminase or serum glutamic-pyruvic transaminase.
51. **'Aspartate transaminase'**, is an enzyme that is found mostly in the liver, but also in muscles. When your liver is damaged, it releases AST into your bloodstream. An AST blood test measures the amount of AST in your blood.
52. **'Gamma-glutamyltransferase'**, an enzyme found in cell membranes of many tissues mainly in the liver, kidney, and pancreas. [1] It is also found in other tissues including intestine, spleen, heart, brain, and seminal vesicles. The highest concentration is in the kidney, but the liver is considered the source of normal enzyme activity. Male and female age 45 years and older: 8-38 units/L or 8-38 international units (IU)/L (SI units)
53. **'Total Bilirubin'**, Normal results for a total bilirubin test are 1.2 milligrams per deciliter (mg/dL) for adults and usually 1 mg/dL for those under 18. Normal results for direct bilirubin are generally 0.3 mg/dL.
54. **'Direct Bilirubin'**, direct bilirubin travels freely through your bloodstream to your liver. Most of this bilirubin passes into the small intestine. A very small amount passes into your kidneys and is excreted in your urine. This bilirubin also gives urine its distinctive yellow color.
55. **'Indirect Bilirubin'**, Indirect bilirubin is the difference between total and direct bilirubin. Common causes of higher indirect bilirubin include: Hemolytic anemia.
56. **'Alkaline phosphatase'**, Alkaline phosphatase (ALP) is an enzyme that is present in many parts of the body, but it is primarily found in the liver, bones, intestine, and kidneys. Alkaline phosphatase testing measures the amount of this enzyme in the blood. One common reference range is from 44 to 147 IU/L.
57. **'Ionized calcium'**, Children: 4.8 to 5.3 milligrams per deciliter (mg/dL) or 1.20 to 1.32 millimoles per liter (millimol/L)
58. Adults: 4.8 to 5.6 mg/dL or 1.20 to 1.40 millimol/L
59. **'Strepto A'**, Streptococcus pyogenes is a species of Gram-positive, aerotolerant bacterium in the genus Streptococcus. These bacteria are extracellular, and made up of non-motile and non-spore-forming cocci. It is clinically important for humans. It is an infrequent, but usually pathogenic, part of the skin microbiota.
60. **'Magnesium'**, normal range for blood magnesium level is 1.7 to 2.2 mg/dL (0.85 to 1.10 mmol/L).
61. **'pCO2 (venous blood gas analysis)'**, is the measure of carbon dioxide within arterial or venous blood. It often serves as a marker of sufficient alveolar ventilation within the lungs. Generally,

under normal physiologic conditions, the value of PCO₂ ranges between 35 to 45 mmHg, or 4.7 to 6.0 kPa

62. **'Hb saturation (venous blood gas analysis)'**, the percentage of total hemoglobin binding sites available for binding to oxygen that is occupied with oxygen.
63. **'Base excess (venous blood gas analysis)'**, is defined as the amount of strong acid that must be added to each liter of fully oxygenated blood to return the pH to 7.40 at a temperature of 37°C and a pCO₂ of 40 mmHg (5.3 kPa).
64. **'pO₂ (venous blood gas analysis)'**, Most healthy adults have a PaO₂ within the normal range of 80–100 mmHg. If a PaO₂ level is lower than 80 mmHg, it means that a person is not getting enough oxygen
65. **'Fio₂ (venous blood gas analysis)'**, FiO₂ (the fraction of inspired oxygen) is defined as the percentage or concentration of oxygen that a person inhales.
66. **'Total CO₂ (venous blood gas analysis)'**, The normal range is 23 to 29 milliequivalents per liter (mEq/L) or 23 to 29 millimoles per liter (mmol/L).
67. **'pH (venous blood gas analysis)'**, Human life requires a tightly controlled pH level in the serum of about 7.4 (a slightly alkaline range of 7.35 to 7.45) to survive .
68. **'HCO₃ (venous blood gas analysis)'**, A bicarbonate level that is higher or lower than normal may mean that the body is having trouble maintaining its acid-base balance, either by failing to remove carbon dioxide through the lungs or the kidneys or perhaps because of an electrolyte imbalance, particularly a deficiency of potassium.
69. **'Rods #'**,
70. **'Segmented'**,
71. **'Promyelocytes'**, A promyelocyte (or progranulocyte) is a granulocyte precursor, developing from the myeloblast and developing into the myelocyte. Promyelocytes measure 12–20 microns in diameter.
72. **'Metamyelocytes'**, A metamyelocyte is a cell undergoing granulopoiesis, derived from a myelocyte, and leading to a band cell. It is characterized by the appearance of a bent nucleus, cytoplasmic granules, and the absence of visible nucleoli
73. **'Myelocytes'**, A myelocyte is a young cell of the granulocytic series, occurring normally in bone marrow (can be found in circulating blood when caused by certain diseases)
74. **'Myeloblasts'**, a unipotent stem cell which differentiates into the effectors of the granulocyte series.
75. **'Urine - Esterase'**, an enzyme present in most white blood cells (WBCs). A few white blood cells are normally present in urine and usually give a negative chemical test result. When the number of WBCs in urine increases significantly, this screening test will become positive. Results of this test will be considered along with a microscopic examination for WBCs in the urine.
76. **'Urine - Aspect'**,

77. 'Urine - pH',
 - a. The acidity/basic of the urine should be between 6.5 and 7.25
78. 'Urine - Hemoglobin',
 - a. Causes dark urine and is a sign of disease
79. 'Urine - Bile pigments',
 - a. An indication of liver dysfunction causes Jaundice
 - b. Character
80. 'Urine - Ketone Bodies',
 - a. Chemical byproduct of lack of glucose(think keto diet)
 - b. character
81. 'Urine - Nitrite',
 - a. Sign of a UTI caused by conversion of Nitrate to Nitrite by bacteria should be Negative
 - b. character
82. 'Urine - Density'
 - a. Relates to kidney function total amount of dissolved substances. Should be between 1.005 and 1.030
 - b. Min -1.3,Max .65 Numeric
83. 'Urine - Urobilinogen'
 - a. Should be a middling amount a lack of it indicates disease <2.0 mg, like cirrhosis
 - b. Character
84. 'Urine - Protein'
 - a. Should be a small amount large means problems < 150mg
 - b. Character
85. 'Urine - Sugar',
 - a. Too much means Diabetes should be 0 to 0.8 mmo/l
 - b. NA
86. 'Urine - Leukocytes',
 - a. Sign of infection to many white blood cells 0-5 WBCs per high power field (wbc/hpf)
 - b. NA
87. 'Urine - Crystals',
 - a. To high might lead to kidney stones <45mg/day
 - b. NA
88. 'Urine - Red blood cells',
 - a. Should be very small < 4 cells per high power field, peeing blood is bad...
 - b. Min = -.203 Max.432
89. 'Urine - Hyaline cylinders',
 - a. 0-5 per LPF means sluggish urine flow
 - b. NA
90. 'Urine - Granular cylinders',
 - a. red blood cells, white blood cells, fat bodies wrapped in a protein matrix 0 -2 per low power field more means disease
 - b. NA

91. 'Urine - Yeasts',
 - a. UTI 1 -15 mg/dL
 - b. NA
92. 'Urine - Color',
 - a. Hydration level - lighter is better roughly 8 pigments
 - b. NA
93. '**Partial thromboplastin time**',
 - a. Time to blood clot should be 25 to 35
 - b. NA
94. '**Relationship (Patient/Normal)**',
 - a. Min 0.813 Max .9057
95. '**International normalized ratio (INR)**',
 - a. Min -.7786 Max .8624
96. '**Lactic Dehydrogenase**'
 - a. Acute/Chronic Tissue damage roughly 140 to 280 U/L
 - b. Min -1.04 Max 2.56
97. '**Prothrombin time (PT), Activity**',
 - a. Time to blood clot 10 to 14 seconds
 - b. NA
98. '**Vitamin B12**',
 - a. 160 to 950 per day to low can cause anemia
 - b. NA
99. '**Creatine phosphokinase**',
 - a. 10 to 120 mcg/L - to high is indicative to a host of diseases
 - b. Min -.46 Max 2.4
100. '**Ferritin**',
 - a. Men: 24 - 336 mcg/L, Women: 11 to 307 mcg/L to low could be anemic
 - b. Min .3959 Max .7763
101. 'Arterial Lactic Acid',
 - a. 0.5 to 1.6 mmol/L high levels could be a host of heart/kidney problems
 - b. Min -.959 Max 3.0041
102. 'Lipase dosage',
 - a. 2,500 lipase units/kg of body weight per meal helps to digest fats
 - b. T/F
103. 'D-Dimer',
 - a. Used to test for bloodclots less than .5 for a negative test
 - b. NA
104. 'Albumin'
 - a. 3.4 to 5.4 g/L too low means liver disease
 - b. NA
105. 'Hb saturation (arterial blood gases)'
 - a. Amount of oxygen in the blood
 - b. NA

106. 'pCO₂ (arterial blood gas analysis)',
 - a. Partial pressure of CO₂ in the blood normal is 35 - 46 mm/Hg
 - b. Min -1.3 to 1.2
107. 'Base excess (arterial blood gas analysis)',
 - a. Min 1.3 Max 1.2245
108. 'pH (arterial blood gas analysis)',
 - a. 7.35 to 7.45 to high dehydration to low is various metabolic issues
 - b. Min -.2249 Max 1.225
109. 'Total CO₂ (arterial blood gas analysis)',
 - a. Amount of CO₂ in the veins
 - b. Min -1.5 Max 1.21
110. 'HCO₃ (arterial blood gas analysis)',
 - a. Amount of anions in blood platelets is 22-26
 - b. Min 1.5468 Max 1.2895
111. 'pO₂ (arterial blood gas analysis)',
 - a. Pressure of Oxygen in blood 80 to 100 mm/Hg
 - b. Min 1.05 Max 2.2
112. 'Arterial Fio₂'
 - a. PaO₂/FiO₂ ratio is the ratio of arterial oxygen partial pressure (PaO₂ in mmHg) to fractional inspired oxygen
 - b. Min 1.5329 Max -.017
113. 'Phosphor'
 - a. 2.8 to 4.5 mg/dL is needed for healthy bones
 - b. -.533 the only value
114. 'ctO₂ (arterial blood gas analysis)'
 - a. Total amount of carbon dioxide in the body
 - b. Min -1.15 Max 1.83