CLAIMS

1. A method for determining a data exception, the method comprising:

receiving a plurality of first data groups that each contain respective pieces of data and that each is to be validated, wherein each of the pieces of data in each first data group corresponds to a same predetermined feature, wherein the pieces of data are encrypted, and wherein the first data groups and the pieces of data are provided by a data-providing platform;

for each first data group in the plurality of first data groups:

obtaining from the first data group a plurality of pieces of data that is to be validated;

obtaining (22) a plurality of pieces of historical data of the data that is to be validated as a comparison data group, wherein the pieces of historical data are provided by the data-providing platform, wherein the data in the first data group and the data in the comparison data group have the same data source, and wherein the comparison data group has no data exceptions;

performing (23) a first two-group significance test on the first data group and the comparison data group, wherein the two-group significance test is a test to determine a probability that (i) a population mean of a population from which the data group is drawn and (ii) a population mean of a population from which the comparison data group is drawn are equal; and

determining (24) whether there is a data exception in the first data group based on a test result of the first two-group significance test, and if there is a data exception, dividing the first data group into a predetermined quantity of second data groups that is to be validated and that each has a smaller size than that of the first data group and performing a third two-group significance test on each second data group, and otherwise accepting the plurality of pieces of data for a next step of processing.

2. The method according to claim 1, wherein obtaining the plurality of pieces of historical data comprises:

obtaining a plurality of groups of historical data of the data that is to be validated, performing a second two-group significance test on every two groups of historical data, and

determining a group of historical data that has no exception as the comparison data group based on a test result of the second two-group significance test.

3. The method according to claim 1, wherein performing the third two-group significance test on each second data group further comprises:

performing the third two-group significance test on each second data group and the comparison data group, and determining a particular second data group has a data exception based on a test result of the third two-group significance test.

4. The method according to any one of claims 1 to 3, wherein the first two-group significance test, the second two-group significance test, and the third two-group significance test are separately one of the following tests: a t test, a z test, and a t' test.

5. The method according to claim 1, further comprising: before performing the first two-group significance test on the first data group that is to be validated and the comparison data group, performing same preprocessing on data in the first data group that is to be validated and data in the comparison data group based on a common distribution feature of the first data group that is to be validated and the comparison data group.

6. The method according to claim 5, wherein the preprocessing is one of the following data transformations: a logarithmic transformation, a square root transformation, a reciprocal transformation, and an arcsine square root transformation.

7. The method according to any one of claims 1 to 6, wherein the predetermined feature defines a type of numerical values within a predetermined range.

8. The method according to any one of claims 1 to 7, further comprising:

setting (31) a hypothesis associated with the plurality of pieces of data;

calculating (32) test statistics to verify hypothesis;

calculating (33) a degree of freedom of the test statistics; and

determining (34) a first probability that the hypothesis is true based on the test statistics and the degree of freedom of the test statistics.

9. The method according to claim 8, further comprising:

comparing the first probability to a critical value to determine a degree of confidence.

10. An apparatus for determining a data exception, the apparatus comprising a plurality of modules configured to perform the method of any one of claims 1 to 9.