Statistical Analysis Using SPSS for Windows — Getting Started —

(Ver. 2026/10/06)

The numbers of figures in the SPSS_screenshot.pptx are shown in red.

1. How to display English messages from IBM SPSS Statistics
(change from Japanese to English)

1) Double click the shortcut icon for IBM SPSS Statistics 32 on the desktop. 0-1
2) From the "編集"(edit) menu, select the "オプション"(options). 0-2
3) In the "オプション"(options) window, click the “言語”(language) tab. You can find the "言語" (language) section in the left-upper corner of the window. 0-3
4) From the "出力"(output) menu, select "英語"(English). From the "ユーザーインタフェイス" (user interface) menu, also select "英語"(English). 0-4
5) Click the "OK" button, then you will see a message window from IBM SPSS Statistics 32. Click the "OK" button in the window.
6) Now you can use SPSS in English!

2. How to restore Japanese messages from IBM SPSS Statistics
(change from English to Japanese)

1) From the "Edit" menu, select the "Options". 0-5
2) In the "Options" window, click the "Language" tab. You can find the "Language" section in the left-upper corner of the window. 0-6
3) From the "Output" menu, select "Japanese". From the "User Interface" menu, also select "Japanese". 0-7
4) Click the "OK" button, then you will see a message window from IBM SPSS Statistics 32. Click the "OK" button in the window.
5) Now you can use SPSS in Japanese!

1. Preparing data for SPSS

Task 1-1 Convert Excel data into SPSS data

1) Preparation: download the Excel file for practice from the following Internet address.
https://www.med.nagoya-u.ac.jp/yobo/spss/dl/data_spss.xls
2) Read the Excel file: from the “File” menu, select the “Open” and then the “Data”. 1-1
3) In the “Open Data” window, change the “Files of type” from SPSS Statistics to Excel. 1-2
4) Double click the “data_spss.xls” file in the central window. 1-3

5) Confirm that the “Are variable names included at the top of your file: Yes” is checked, and click the “OK” button. 1-4
6) The Excel data has now been converted into SPSS data.
7) Save the new SPSS data. From the “File” menu, select the “Save as”. You can input the new file name and save the SPSS data file. 1-5, 1-6

Task 1-2 Label variables and their values

1) Click the “Variable View” tab at the left-lower corner of the data editor window.
2) You can label a variable by inputting any label in the “Label” field or cell. 1-7
3) In addition, you can tag values of a variable by clicking the right end of the “Values” field or cell. In the “Value Labels” window, click ⊕ button and input a value and its corresponding label. Repeat this procedure until you label all the values (ex. “men” for 1 and “women” for 2 in the variable “sex”) and click the “OK” button. 1-8, 1-9, 1-10

2. Converting variables

Task 2-1 Create a new variable, bmi, from height and weight.

1) From the “Transform” menu, select the “Compute Variable”. 2-1
2) Input “bmi” in the “Target Variable” box and “wt/hei/hei*10000” in the “Numeric Expression” box. 2-2
3) Click “OK” button at the bottom of the window.

Task 2-2 Categorize BMI into three groups: < 18.5, 18.5-24.9, 25.0+ (kg/m^2).

1) From the “Transform” menu, select the “Recode into Different Variables”. 2-3
2) Select the variable “bmi” and click the → button. 2-4
3) In the “Output Variable” section, input “bmi3g” in the “Name” box and “BMI 3 groups” in the “Label” box. 2-5
4) Click the “Change” button.
5) Click the “Old and New Values” button at the center of the window. 2-6
   a) Select the “Range, value through HIGHEST:” button and input “25” in the box. Input “3” in the “Value” box in the “New Value” section and click the “Add” button. 2-7
   b) Select the “Range:” button. Input “18.5” in the “Range:” box and “25” in the “through” box. Input “2” in the “Value” box and click the “Add” button. 2-8
   c) Select the “Range, LOWEST through value:” button. Input “18.5” in the box. Input “1” in the “Value” box and click the “Add” button. 2-9
   d) Click the “Continue” button at the bottom of the window.
   e) Click the “OK” button in the “Recode into Different Variables” window. 2-10

- We recommend you to label the three values for the variable “bmi3g” (e.g., lean, normal,

overweight).

3. Statistical analysis

Task 3-1 Frequency tables for sex and age

1) From the “Analyze” menu, select the “Descriptive Statistics” and then the “Frequencies”. 3-1
2) Move the variables sex and age into the “Variable(s)” box. (Select target variables and click the ⟶ button. 3-2
3) Click the “OK” button at the left-lower corner of the window. 3-3

Task 3-2 Descriptive Statistics (means, SD, maximums and minimums) for age, height, weight, systolic blood pressure, and diastolic blood pressure

1) From the “Analyze” menu, select the “Descriptive Statistics” and then the “Descriptives”. 3-4
2) Move the five variables (age, hei, wt, sbp, and dbp) into the “Variable(s)” box. (Select target variables and click the ⟶ button. 3-5
3) Click the “OK” button at the left-lower corner of the window. 3-6

Task 3-3 Correlation coefficients among age, BMI, systolic blood pressure, and diastolic blood pressure

1) From the “Analyze” menu, select the “Correlate” and then the “Bivariate with Confidence Interval. 3-7
2) Move the four variables (age, bmi, sbp, and dbp) into the “Variables” box. (Select target variables and click the ⟶ button. 3-8
3) Click the “OK” button at the left-lower corner of the window. 3-9

Task 3-4 Scatter graph to depict the association between BMI and systolic blood pressure

1) From the “Graphs” menu, select the “Scatter/Dot”. 3-10
2) Select the “Simple Scatter” and click the “Define” button. 3-11
4) Move the variables sbp and bmi into the “Y Axis” box and “X axis” one, respectively, and click the “OK” button at the left-lower corner of the window. 3-12, 3-13

Task 3-5 Mean systolic blood pressure by the BMI category (< 18.5, 18.5-24.9, 25.0+ [kg/m^2])

1) From the “Analyze” menu, select the “Compare Means and Proportions” and then the “Means”. 3-14
2) Move the variables sbp and bmi3g into the “Dependent List” box and “Independent List” one, respectively, and click the “OK” button at the left-lower corner of the window. 3-15, 3-16

- You can add other statistics by clicking “Options” button at the right-upper corner of the “Means” window.

Task 3-6 Statistical test for the difference in systolic blood pressure between sexes (unpaired *t* test)

1) From the "Analyze" menu, select the "Compare Means and Proportions" and then the "Independent-Samples T Test". 3-17
2) Move the variables sbp and sex into the "Test Variable(s)" box and "Grouping Variable" one, respectively. 3-18
3) Click the "Define Groups" button under the "Grouping Variable" box. Input "1" in the "Group 1" box and "2" in the "Group 2" box, and then click the "Continue". 3-19
4) Check the "Homogeneity of variance test" checkbox and click the "OK" button at the left-lower corner of the "Independent-Samples T Test" window. 3-20

Task 3-7 Statistical test for the association between the BMI category (< 18.5, 18.5-24.9, 25.0+ [kg/m^2]) and sex (cross tables and chi-square tests)

1) From the "Analyze" menu, select the "Descriptive Statistics" and then the "Crosstabs". 3-21
2) Move the variables sex and bmi3g into the "Row(s)" box and "Column(s)" one, respectively. 3-22
3) Click the "Statistics" button at the right-upper corner of the window. Check the "Chi-square" checkbox, and then click the "Continue". 3-23
4) Click the "OK" button at the left-lower corner of the "Crosstabs" window.

- You can add percentages by clicking "Cells" button at the right-upper corner of the "Crosstabs" window. Check the "Row", "Column" and/or "Total" checkboxes in the "Percentages" section and click the "Continue" button.

Task 3-8 Statistical test for the difference in systolic blood pressure among the three BMI categories (< 18.5, 18.5-24.9, 25.0+ [kg/m^2]; one-way analysis of variance [one-way ANOVA])

Method A:

1) From the "Analyze" menu, select the "Compare Means and Proportions" and then the "One-Way ANOVA". 3-24
2) Move the variables sbp and bmi3g into the "Dependent List" box and "Factor" one, respectively. 3-25
3) Click the "Post Hoc" button at the right-upper corner of the window.
4) Check the "Tukey" checkbox and click the "Continue". 3-26

- You can use other methods for multiple comparisons by checking the corresponding checkboxes.

5) Click the "OK" button at the left-lower corner of the "One-Way ANOVA" window. 3-27

Method B:

1) From the "Analyze" menu, select the "General Linear Model" and then the "Univariate". 3-28

2) Move the variables sbp and bmi3g into the “Dependent Variable” box and “Fixed Factor(s)” one, respectively. 3-29
3) Click the “OK” button at the left-lower corner of the “Univariate” window.

- You can use methods for multiple comparisons by clicking the “Post Hoc” button in the “Univariate” window.

Task 3-9 Establish a linear regression model that predicts systolic blood pressure from sex, age, and BMI.

1) From the “Analyze” menu, select the “Regression” and then the “Linear”. 3-30
2) Move the variable sbp into the “Dependent” box and put the variables sex, age, and bmi into the “Independent(s)” box. 3-31
3) Click the “OK” button at the left-lower corner of the window.

Exercises

Exercise 1 Categorize age into four groups: 30-39, 40-49, 50-59, and 60-69 years. Then, compute the mean of systolic blood pressure by the age category.

Exercise 2 Establish a linear regression model that predicts mean blood pressure from three variables, that is, sex, age, and BMI. Compute mean blood pressure using the formula below.

mean blood pressure = ([systolic blood pressure] + 2×[diastolic blood pressure])/3

Appendix: Logistic regression analysis

You can conduct logistic regression analysis with SPSS by selecting the “Regression” and then the “Binary Logistic” from the “Analyze” menu.

Ex. Examine the association of hypertension with sex, age, and BMI by multiple logistic regression analysis. Define hypertension as systolic blood pressure of 140+ mmHg and/or diastolic blood pressure of 90+ mmHg. Categorize age into four groups: 30-39, 40-49, 50-59, and 60-69 years.

1) Create a new variable, hyt, from sbp and dbp using the following equation (ref. Task 2-1).
   hyt = sbp >=140 | dbp >=90 (“|” means “or”)
2) Categorize age into four groups: 30-39, 40-49, 50-59, and 60-69 years (ref. Task 2-2 and Exercise 1). In this example, the variable for these groups is named “age4g”.
3) From the “Analyze” menu, select the “Regression” and then the “Binary Logistic”.
4) Move the variable hyt into the “Dependent” box and put the variables sex, bmi, and age4g into

the “Covariates” box.

5) Click the “Categorical” button at the right-upper corner of the window.
6) Move the variable age4g into the “Categorical Covariates” box. Click the “First” button in the “Change Contrast” section.
7) Click the “Continue” button at the bottom of the window.
8) Click the “Options” button at the right-upper corner of the window.
9) Check the “CI for exp(B):” checkbox in the “Statistics and Plots” section.
10) Click “Continue” in the “Logistic Regression: Options” window.
11) Click the “OK” button at the left-lower corner of the “Logistic Regression” window.