JASP for Audit
User Manual

Statistical Auditing Group

"The best things in life are free."

Contents

Contents

Preface

The **JASP for Audit User Manual** provides detailed instructions and best practices for working with the Audit module in the free and open-source software JASP. It covers various aspects, including data import, analysis techniques, and interpretation of results. The manual is curated by the Statistical Auditing Group at Nyenrode Business University, ensuring that users have access to reliable and up-to-date information.

Getting Started

Statistical theory is fundamental to many auditing procedures. To perform these procedures effectively, auditors need user-friendly software for statistical analyses and the knowledge to interpret the results. JASP (JASP Team, 2025) is an open-source, free-of-charge, cross-platform statistical software program that supports statistical auditing through its Audit module (Derks et al., 2021).

The Audit module (i.e., JASP for Audit) allows auditors to plan, execute, and interpret a wide range of statistical auditing procedures using state-of-the-art statistical methods, thereby reducing programming errors and simplifying the process. Tailored for auditors, the module features an intuitive interface that aligns with audit processes and international standards on auditing. In addition to standard frequentist methods, the Audit module incorporates Bayesian methods to enhance audit transparency and efficiency by utilizing existing information.

In summary, the Audit module takes care of the complex statistical work, allowing you to concentrate on interpreting the results of your analysis. The remaining paragraphs in this chapter discuss how to get started using JASP for Audit.

Downloading JASP

JASP for Audit is part of JASP, which can be freely downloaded from www.jasp-stats.org. Click the 'Download JASP' button on the homepage to access the download page and choose your preferred installation. JASP is available for Windows, MacOS, Linux, and Chrome OS.
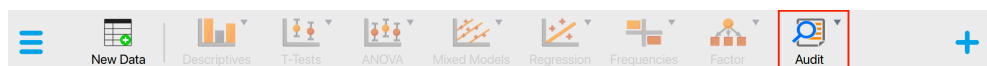
Enabling the Audit module

After opening JASP, you will see the following main menu bar at the top of the screen.
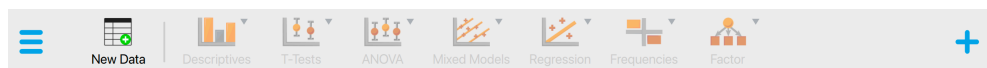
To find the Audit module, click the '+' icon on the right of this menu bar. A different menu will appear on the right side which shows all available modules. Check the box next to 'Audit' to make the module visible in the main menu bar. You can now access the Audit module and its analyses by clicking its module icon in the menu bar (see image below).


Accessibility

The Audit module is a robust tool for statistical auditing. The following paragraphs detail its accessibility features, including where to locate help files and how to the reliability of the statistical results is ensured.

Help files

Once you open an analysis in the Audit module, you can click the blue 'i' icon next to the analysis title to access a help file that explains its functionality. Additional help files for certain options can be accessed by clicking the blue 'i' icon next to those options.
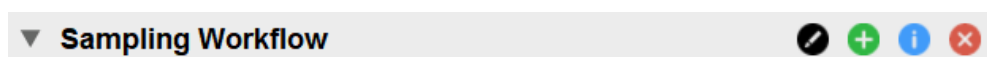
Validation of statistical results

The statistical results generated by the Audit module are based on the R package jfa (Derks, 2025). For comprehensive documentation and information on the benchmarks used for validation, please visit the package website at https://koenderks.github.io/jfa/.

Part I

Audit Sampling

Chapter 1

Sampling Workflow

The goal of statistical audit sampling is to infer the misstatement in a population based on a representative sample. This can be challenging, but the Audit module simplifies the process into four stages: planning, selection, execution, and evaluation.

| Planning | Selection | Execution | Evaluation |

More detailed information about the individual stages in the audit sampling workflow is provided below.

1.1 The four stages of the sampling workflow

In the planning stage, you determine the sample size needed to support the assertion that the population's misstatement is below the performance materiality. This involves using prior audit outcomes and information about inherent risk and control risk. Expectations about error rates also influence the sample size required to maintain statistical confidence.

Using the sample size from the planning stage, you select a statistically representative sample. Each sampling unit receives an inclusion probability, and units are selected based on these probabilities. Monetary unit sampling assigns probabilities to individual monetary units, making higher-value items more likely to be selected. Record sampling assigns equal probabilities to all items.

In the execution stage, you assess the correctness of selected items. The simplest method categorizes items as correct or incorrect, while a more accurate method considers the true value (audit value) of items. Annotating samples with audit values provides a more precise estimate of misstatement. If book values are unavailable, use the correct/incorrect method.

In the evaluation stage, you use the annotated sample to infer the total misstatement

in the population. Statistical techniques calculate a projected maximum misstatement, and the population is approved if this is below the performance materiality.

This manual emphasizes the practical application of the audit sampling workflow in JASP. For a deeper understanding of the statistical theory behind the four stages of the audit sampling workflow, read the free online book Statistical Audit Sampling with R.

1.2   A practical example in JASP

The Audit module in JASP offers two ways to navigate the audit sampling workflow: the Sampling Workflow analysis, which guides you through all four stages, and individual analyses for Planning, Selection, and Evaluation. This chapter uses the classical sampling workflow analysis to explain the Audit module's core functionality. Note that a Bayesian variant of the sampling workflow is also available.

Let's explore an example of the audit sampling workflow. To follow along, open the 'Testing for Overstatements' dataset from the Data Library. Navigate to the top-left menu, click 'Open', then 'Data Library', select '7. Audit', and finally click on the text 'Testing for Overstatements' (not the green JASP-icon button).
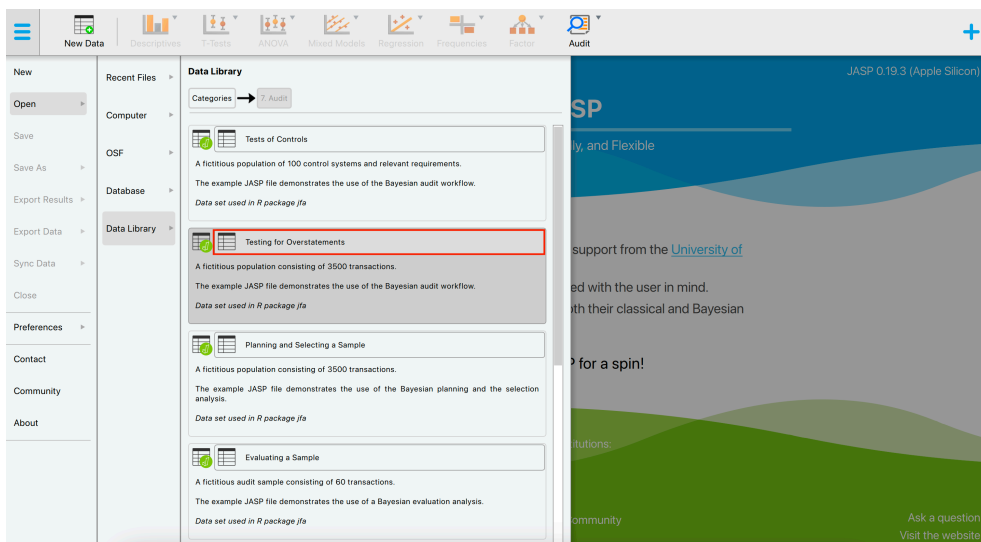
This will open a dataset with three columns: 'ID', 'bookValue', and 'auditValue'. The 'ID' column represents the identification number of the items in the population. The 'bookValue' column shows the recorded values of the items, while the 'auditValue' column displays the true values. The 'auditValue' column is included for illustrative purposes, as auditors typically know the true values only for the audited sample, not for all items in the population.

1.2.1 Stage 1: Planning

To start the sampling workflow, click on the Audit module icon and select 'Sampling Workflow'. This will open the following interface, where you need to fill in the options for the statistical analysis.

The following steps are required:

1. **Indicate the variables:** First, enter the variable indicating the identification numbers of the items in the corresponding box. Optionally, if you have access to the book values of the items, you can enter this variable as well.

2. **Sampling objectives:** Next, formulate your sampling objectives. Enable the 'Performance materiality' objective if you want to test whether the total misstatement in the population exceeds a certain limit (i.e., the performance materiality). This approach allows you to plan a sample such that, when the sample meets your expectations, the maximum error is said to be below performance materiality. Enable the 'Minimum precision' objective if you want to obtain a required minimum precision when estimating the total misstatement in the population. This approach allows you to plan a sample such that, when the sample meets expectations, the uncertainty of your estimate is within a tolerable percentage. In the example, we choose a performance materiality of 3.5%.

3. **Expected misstatement:** Then, indicate how many misstatements are tolerable in the sample. In the example, we choose to tolerate one full misstatement in the sample.

4. **Prior information:** Additionally, indicate the risks of material misstatement via the audit risk model. According to the Audit Risk Model, audit risk can be divided into three constituents: inherent risk, control risk, and detection risk. Inherent risk is the risk posed by an error in a financial statement due to a factor other than a failure of internal controls. Control risk is the probability that a material misstatement is not prevented or detected by the internal control systems of the company (e.g., computer-managed databases). Both these risks are commonly assessed by the auditor on a 3-point scale consisting of low, medium, and high. Detection risk is the probability that an auditor will fail to find material misstatements in an organization's financial statements. For a given level of audit risk, the tolerable level of detection risk bears an inverse relationship to the other two assessed risks. Intuitively, a greater risk of material misstatement should require a lower tolerable detection risk and, accordingly, more persuasive audit evidence. In this example, we choose to set all risks to 'High' and solely rely on evidence from substantive testing.

The primary output from the planning stage, shown below, indicates that a minimum sample size of 134 sampling units is required to achieve 95% assurance that the misstatement in the population is below 3.5%, while allowing for one misstatement in the sample.

Table 1. *Planning Summary*

	Value
Performance materiality	0.035
Inherent risk	1.000
Control risk	1.000
Analytical risk	1.000
Detection risk	0.050
Tolerable misstatements	1.000
Minimum sample size [a]	134

Note. The minimum sample size is based on the binomial distribution (p = 0.035)
[a] Based on this sample size, the selection interval spans 10471.8 units.

5. **Next stage:** Finally, progress to the selection stage by clicking the 'To Selection' button.

For a more detailed explanation of the options and output in the planning stage, see Chapter 2.

1.2.2 Stage 2: Selection

In the selection stage, you must select the 134 sampling units from the population. Once the 'To Selection' button is pressed, the interface from the selection stage opens.

The following steps are required:

1. **Randomness options:** Begin by selecting the options related to randomness in the selection procedure. The seed option is important as it ensures that random procedures are reproducible, allowing for consistent results across multiple runs. A random number will be chosen each time you start the analysis. Additionally, the 'Randomize item order' option is available to randomly shuffle the rows in the dataset, which can help mitigate any biases that might arise from the original order of the data.

2. **Sampling units:** Next, specify the sampling units for the selection process. These units can either be items or monetary units. If no book value variable

is provided, the sampling units default to 'Items', enabling attribute sampling. Conversely, if a book value variable was indicated during the planning stage, the sampling units default to 'Monetary units', facilitating monetary unit sampling (MUS). MUS is particularly useful for auditing financial data as it considers the monetary value of each unit.

3. **Sampling method:** Then, choose the selection method to be used in the sampling process. The available algorithms include:

   - **Fixed interval sampling:** This method selects units at regular intervals from the dataset, ensuring a systematic sampling approach.
   - **Cell sampling:** This technique involves dividing the dataset into cells and randomly selecting units from each cell, promoting a systematic sampling approach with a bit of randomness.
   - **Random sampling:** This approach randomly selects units from the entire dataset, providing a simple yet effective method for ensuring randomness.

The primary output from the selection stage, as shown in the first table below, reveals that 134 sampling units were selected from 134 items. The sample's total value amounts to €67,821.22, representing 4.8% of the total population value. The second table provides details specific to interval selection using monetary unit sampling. It indicates the number of items selected in the 'Top stratum', which includes all items larger than a single interval (for fixed interval selection). In this instance, there were 0 items in the top stratum.

Table 3. Selection Summary ▼

No. units	No. items	Selection value	% of population value
134	134	€67,821.22	4.8%

Note. From each of the intervals of size 10471.8, unit 9584 is selected using seed 300.

Table 4. Information about Monetary Interval Selection

	Items	Value	Selected items	Selected units	Selection value	% of total value
Total	3,500	€1,403,220.82	134	134	€67,821.22	4.8%
Top stratum	0	€0	0	0	€0	0%
Bottom stratum	3,500	€1,403,220.82	134	134	€67,821.22	4.8%

Note. The top stratum consists of all items with a book value larger than a single interval.

4. **Next stage:** Finally, progress to the execution stage by clicking the 'To Execution' button.

1.2.3 Stage 3: Execution

In the execution stage, you must judge the fairness of the 134 sampled items. Once the 'To Execution' button is pressed, the interface from the execution stage opens.

The following steps are required:

1. **Annotation method:** First, decide how to annotate the selected items. You have two options:

   - Audit value: Annotate the items with their audit (true) values. This method is recommended (and automatically selected) when the items have a monetary value.
   - Correct / Incorrect: Annotate the items as correct (0) or incorrect (1). This method is recommended (and automatically selected) when the items do not have a monetary value.

2. **Column names:** Next, specify the names of the two columns that will be added to the dataset. The first column name will indicate the result of the selection, while the second column name will contain the annotation of the items. Click the 'Continue' button to confirm the options and open the data viewer.

3. **Annotating items:** Then, use the data viewer to annotate the selected items with their book value. For example, in this case, item 50826 (row 25, highlighted in red) had a book value of €333.03 but a true value of €200. The remaining items have correctly reported book values.

4. **Next stage:** Finally, progress to the evaluation stage by clicking the 'To Evaluation' button.

1.2.4 Stage 4: Evaluation

In the evaluation stage, you assess the misstatement in the sample and extrapolate it to the entire population. Once you press the 'To Evaluation' button, the interface for the evaluation stage will open.

The following steps are required:

1. **Annotation variable:** Specify the variable that contains the annotation of the items in the corresponding box.

2. **Additional tablaes:** It is recommended to request the 'Misstated items' table from the 'Report' section. This table displays the items in the sample where the book value did not match the true value. Additional tables and figures to clarify the output, which will be discussed in Chapter 4, can be requested here as well.

The primary output from the evaluation stage, as shown in the first table below, indicates that the most likely misstatement in the population is estimated to be 0.003, or 0.3%. The 95% upper bound for this estimate is 0.027, or 2.7%. This upper bound is lower than the performance materiality of 3.5%, meaning the auditor has achieved at least 95% assurance that the population misstatement is below the performance materiality.

Table 4. *Evaluation Summary*

	Value
Performance materiality	0.035
Sample size	134
Misstatements	1
Taint	0.396
Most likely misstatement	0.003
95% Upper bound	0.027
Precision	0.025
p–value	0.019

Note. The results are computed using the binomial distribution.

Table 5. *Misstated Items*

ID	Book value	Audit value	Difference	Taint	Counted
50,826	€331.03	€200	€131.03	0.396	x1
Total			€131.03	0.396	

Based on the results of this statistical analysis, the auditor concludes that the population is free of material misstatement.

Chapter 2

Planning

Chapter 3

Selection

Chapter 4

# Evaluation

Part II

Data Auditing

Chapter 5

Benford's Law

Benford's law states that the distribution of leading digits in a population naturally follows a certain distribution. In auditing, assessing whether a distribution of digits in the population conforms to Benford's law may provide additional evidence that the transactions in the population might need further investigation. Non-conformity to Benford's law does not necessarily indicate fraud. A Benford's law analysis should therefore only be used to acquire insight into whether a population might need further investigation.

5.1   A practical example in JASP

Let's explore an example analysis of Benford's law. To follow along, open the 'Assessing Benford's Law' dataset from the Data Library. Navigate to the top-left menu, click 'Open', then 'Data Library', select '7. Audit', and finally click on the text 'Assessing Benford's Law' (not the green JASP-icon button).
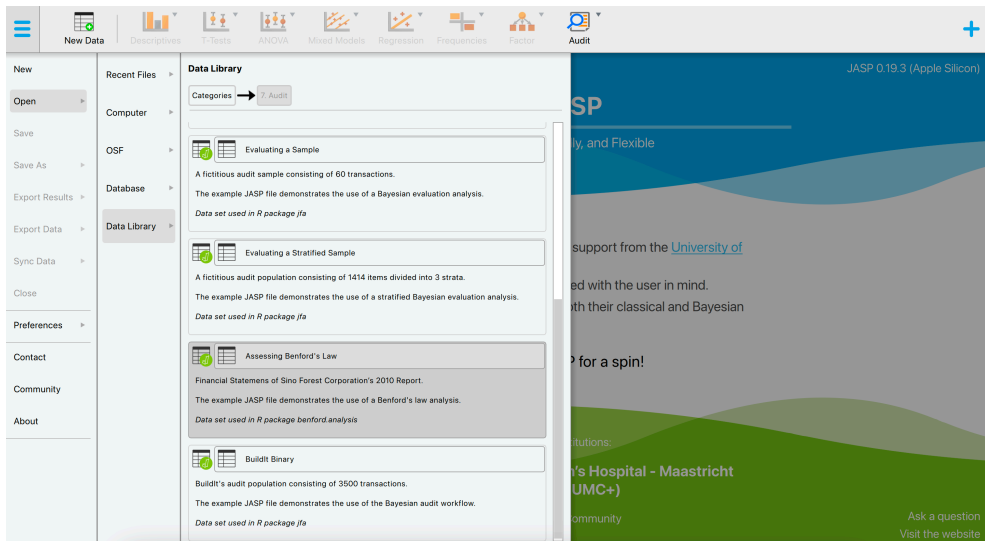
The interface...

The first table in the output...

Table 1. *Omnibus Test – Benford's Law*

	n	MAD	X^2	df	p	BF_{10}[a]
value	772	0.007	7.652	8	0.468	1.449×10^{-7}

Note. The null hypothesis specifies that the first digits (1 – 9) in the data set are distributed according to Benford's law.

[a] The Bayes factor is computed using a Dirichlet($\alpha_1,...,\alpha_9$) prior with $\alpha = 1$.

The second table in the output...

Table 2. *Frequency Table*

Leading digit	Count	Benford's law	Relative frequency	p^a	BF_{10}^b
1	231	0.301	0.299	0.937	0.042
2	124	0.176	0.161	0.277	0.064
3	97	0.125	0.126	0.957	0.030
4	70	0.097	0.091	0.626	0.031
5	64	0.079	0.083	0.689	0.027
6	54	0.067	0.070	0.719	0.024
7	40	0.058	0.052	0.537	0.026
8	54	0.051	0.070	0.022	0.290
9	38	0.046	0.049	0.606	0.022

Note. The null hypothesis specifies that the relative frequency of a digit is equal to its expected relative frequency under Benford's law.
[a] The *p*–values are based on independent binomial distributions.
[b] Bayes factors are computed using a beta(1, 1) prior.

The figure in the output...
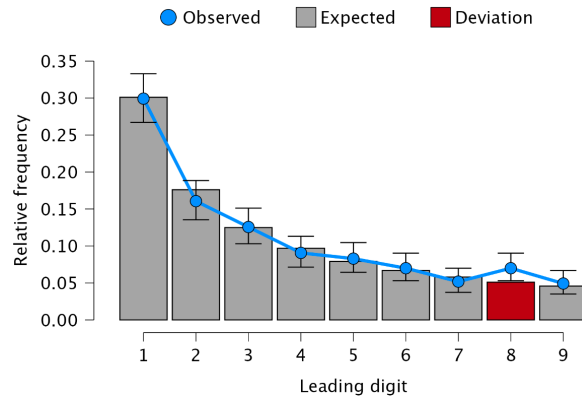
Chapter 6

Repeated Values

The repeated values analysis analyzes the frequency with which values get repeated within a dataset (called "number-bunching") to statistically identify whether the data were likely tampered with. Unlike Benford's law this approach examines the entire number at once, not only the first or last digit (Simohnsohn, 2019).

To determine whether the data show an excessive amount of bunching, the null hypothesis that the data do not contain an unexpected amount of repeated values is tested. To quantify what is expected, this test requires the assumption that the integer portions of the numbers are not associated with their decimal portions.

6.1 A practical example in JASP

Part III

Algorithm Auditing

Chapter 7

Fairness

This is the user manual for **JASP for Audit**, which is a module for the free and open-source statistical software program **JASP** that integrates the functionality of the **jfa** package and offers a user-friendly graphical interface that caters specifically to statistical auditing (**https://jasp-stats.org**).