

AMERICAN HOUSING SURVEY: HOUSING AFFORDABILITY DATA SYSTEM

Download: [Excel Workbook](#), [ReadMe](#)

Abstract: is a set of files derived from the 1985 and later national American Housing Survey (AHS) and the 2002 and later Metro AHS. This system categorizes housing units by affordability and households by income, with respect to the Adjusted Median Income, Fair Market Rent (FMR), and poverty income. It also includes housing cost burden for owner and renter households. These files have been the basis for the worst case needs tables since 2001.¹

Data Set Characteristics:	Multivariate, Time-Series	Number of Instances:	N/A	Area:	Housing
Attribute Characteristics:	N/A	Number of Attributes:	2	Missing Values?	N/A

Source:

<https://www.huduser.gov/portal/datasets/hads/hads.html>

Updated Codebook:

Original Codebook credited to <https://www.huduser.gov/portal/datasets/hads/hads.html>

Updates made by Kim Kirk for Business Statistics and Analysis for Rice University.

¹ <https://www.huduser.gov/portal/datasets/hads/hads.html>

Data Set Information:

The *Housing Affordability Data System (HADS)* is a set of files derived from the 1985 and later national American Housing Survey (AHS) and the 2002 and later Metro AHS. This system categorizes housing units by affordability and households by income, with respect to the Adjusted Median Income, Fair Market Rent (FMR), and poverty income. It also includes housing cost burden for owner and renter households. These files have been the basis for the worst case needs tables since 2001. The data files are available for public use, since they were derived from AHS public use files and the published income limits and FMRs. We are providing these files give the community of housing analysts the opportunity to use a consistent set of affordability measures.²

Data variables include:

Name	Data Type	Explanation
STATUS	Character '1', '3'	Occupied or Vacant
VALUE	Numerical (\$)	Current market value of unit

Data Summaries Calculated:

- Mean
- Standard Error
- Median
- Mode
- Standard Deviation
- Sample Variance
- Kurtosis

² <https://www.huduser.gov/portal/datasets/hads/hads.html>

- Skewness
- Range
- Minimum
- Maximum
- Sum
- Count

Units for Data Summaries and Variables:

All units are dollars (\$)

Data Transformations:

- Dataset was downloaded from host website: <https://www.huduser.gov/portal/datasets/hads/hads.html>
- VALUE and STATUS variable were used for analysis.
- VALUE was determined to be the dependent variable, STATUS was determined to be the independent variable.
- VALUE variable was grouped into categories "unoccupied" and "vacant", reflecting the STATUS of each unit.
- Worksheets were created to hold four categories for the analysis:
- Summary Report identifies and answers the business question.
- Descriptive Statistics identifies and calculates descriptive statistics
- Graphs displays histograms and boxplots
- Test holds results for statistical tests run
- VALUE is numeric continuous, normally distributed. STATUS is nominal, categorical consisting of two independent groups of "occupied" and "vacant".
- Hypothesis tested was if there is a difference between the market VALUE for "occupied" and "vacant" STATUS. A two sample t-Test assuming unequal variances was run to determine statistical significance between group means; results are listed in the "Summary Report" worksheet.