

# SAS: Display IQR in PROC MEANS

Authored by  
**Mohammed looti**

November 15, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *SAS: Display IQR in PROC MEANS*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=1946>

## Introduction to Comprehensive Descriptive Statistics using PROC MEANS

In the rigorous world of statistical analysis, obtaining a concise yet comprehensive summary of your raw data is the foundational first step. The [SAS](#) System, a leading platform for data management and advanced analytics, provides powerful tools for this purpose. Among these tools, the [PROC MEANS](#) procedure stands out as the workhorse for calculating essential [descriptive statistics](#). This versatile procedure allows data analysts to quickly generate crucial measures of central tendency, location, and variability for numeric variables, offering immediate insight into the data's overall distribution and characteristics.

While [PROC MEANS](#) automatically computes common statistics like the mean, minimum, maximum, and standard deviation, it notably omits certain robust measures by default. One of the most important measures often missing from the standard output is the [Interquartile Range \(IQR\)](#). The [IQR](#) is a critical indicator of statistical dispersion, providing a reliable alternative to the standard deviation, particularly when datasets exhibit asymmetry, non-normal distributions, or contain extreme values known as [outliers](#).

Fortunately, integrating the [IQR](#) into your standard [PROC MEANS](#) output is a straightforward process requiring only a specific keyword addition. This article serves as an expert guide, detailing the exact syntax and practical steps required to explicitly request and display the [IQR](#). By the end of this tutorial, you will be equipped to generate more robust and informative data summaries using [SAS](#), thereby ensuring your statistical interpretation is based on the most appropriate measures of variability.

### The Interquartile Range (IQR): A Robust Measure of Dispersion

To properly leverage the [PROC MEANS](#) procedure, it is essential to first understand the metric we are seeking to calculate. The [Interquartile Range \(IQR\)](#) is formally defined as the measure of statistical dispersion that describes the range occupied by the middle 50% of the data points. It is mathematically calculated by taking the difference between the third quartile (Q3, the [75th percentile](#)) and the first quartile (Q1, the [25th percentile](#)). This focus on the central data avoids the distorting influence of extreme values.

The primary strength of employing the [IQR](#) lies in its intrinsic resistance to [outliers](#) and skewness. Standard measures of spread, such as the [standard deviation](#), rely heavily on every data point and are therefore highly susceptible to distortion if a dataset contains even a small number of unusually high or low values. When dealing with real-world datasets--which are rarely perfectly normal--the [IQR](#) provides a more accurate and stable representation of the typical variability found within the core of the distribution.

Furthermore, the [IQR](#) serves as a direct link between tabular [descriptive statistics](#) and graphical

data visualization. Specifically, the length of the box in a [box plot](#) is defined precisely by the [IQR](#). Analysts use this measure not only to understand data spread but also as a fundamental component in common methods for identifying potential [outliers](#) (typically values falling more than 1.5 times the IQR above Q3 or below Q1). Understanding the calculation and role of the [IQR](#) is thus indispensable for comprehensive data exploration and robust reporting.

## Essential Syntax: Requesting IQR via the QRANGE Option

To calculate and output the [IQR](#) using [PROC MEANS](#) in [SAS](#), you must utilize the `QRANGE` keyword within the procedure statement. This keyword acts as an instruction to the [SAS](#) system, compelling it to compute and include the Interquartile Range alongside any default or explicitly requested statistics. This simple addition is key to customizing your summary output for deeper distributional analysis.

The standard syntax for integrating the [IQR](#) requires listing `QRANGE` among the statistic keywords immediately following the `DATA=` option. The following template demonstrates how to request a mix of essential statistics, including the measures of location (Mean), dispersion (Std, QRANGE), and range (Min, Max):

```
proc means data=my_data N Mean QRANGE Std Min Max;  
var points;  
run;
```

This code snippet requests a set of robust [descriptive statistics](#) for the variable `points` within the dataset `my_data`. Below is a detailed explanation of the critical options included in this statement:

`N`: The total count of valid, non-missing observations for the variable, indicating the sample size used in the computation.

`Mean`: Calculates the arithmetic average of the variable, a core measure of central tendency.

`QRANGE`: The mandatory keyword that computes and outputs the [Interquartile Range](#).

`Std`: Provides the [standard deviation](#), useful for comparison against the IQR, especially when assessing data normality.

`Min` and `Max`: Display the lowest and highest values, respectively, defining the overall range of the dataset.

By explicitly listing these statistics, you gain precise control over your summary table, ensuring that the robust measure of spread, the [IQR](#), is always available for informed data interpretation.

## Step-by-Step Example: Calculating IQR for Real-World Data

To illustrate the functionality of the `QRANGE` option, we will utilize a practical example involving a small dataset containing basketball performance metrics. This type of performance data is often prone to high variability and potential [outliers](#) (e.g., a single player having an exceptionally high-scoring game), making the [IQR](#) a particularly valuable statistic for analysis. We will first set up the data within the [SAS](#) environment.

The following code snippet creates a SAS dataset named `my_data`, including variables for `team` (a character variable), and two numeric variables: `points` and `assists`. This creation step ensures we have a valid data structure ready for analysis using [PROC MEANS](#).

```
/*create dataset*/  
data my_data;  
input team $ points assists;  
datalines;  
A 10 2  
A 17 5  
A 17 6  
A 18 3  
A 15 0  
B 10 2  
B 14 5  
B 13 4  
B 29 0  
B 25 2  
C 12 1  
C 30 1  
C 34 3  
C 12 4  
C 11 7  
;  
run;
```

```
/*view dataset*/  
proc print data=my_data;
```

Running the code above generates the dataset structure, which can be visually confirmed using `PROC PRINT`. Visualizing the data ensures that all 15 observations are correctly loaded and that the numeric variables (`points` and `assists`) are correctly formatted for statistical computation.

Obs	team	points	assists
1	A	10	2
2	A	17	5
3	A	17	6
4	A	18	3
5	A	15	0
6	B	10	2
7	B	14	5
8	B	13	4
9	B	29	0
10	B	25	2
11	C	12	1
12	C	30	1
13	C	34	3
14	C	12	4
15	C	11	7

Before we introduce the `QRANGE` option, let's run a basic **PROC MEANS** statement using the default settings for the `points` variable. This step highlights why explicit inclusion of specific statistics is necessary for a complete analysis.

```
/*calculate summary statistics for points variable (Default Output)*/  
proc means data=my_data;  
var points;  
run;
```

The resulting default output, as displayed below, includes the standard set of **descriptive statistics**: the number of observations (N=15), the Mean (19.07), the **Standard Deviation** (8.22), the Minimum (10), and the Maximum (34). Crucially, the Interquartile Range is absent, demonstrating the need to modify the procedure statement to obtain this valuable measure of central spread.

### The MEANS Procedure

Analysis Variable : points				
N	Mean	Std Dev	Minimum	Maximum
15	17.8000000	7.8848861	10.0000000	34.0000000

## Displaying IQR with the QRANGE Option

To rectify the omission in the default output and successfully include the [IQR](#) in our summary table, we must re-run the [PROC MEANS](#) procedure, this time explicitly listing the statistics we require, including the essential [QRANGE](#) keyword. This customization ensures that our analysis benefits from the robust perspective offered by the Interquartile Range.

The updated [SAS](#) code below requests the full set of statistics: N, Mean, [QRANGE](#), Standard Deviation, Minimum, and Maximum. Notice how the statistics are listed directly after the dataset reference:

```
/*calculate summary statistics for points and include IQR*/
proc means data=my_data N Mean QRANGE Std Min Max;
var points;
run;
```

Executing this revised code generates a comprehensive output table that now successfully features the Interquartile Range. This addition fundamentally changes the depth of the statistical summary, providing a direct measure of spread that is not unduly influenced by the extreme scoring values (like the 34 points scored by one player).

### The MEANS Procedure

Analysis Variable : points					
N	Mean	Quartile Range	Std Dev	Minimum	Maximum
15	17.8000000	13.0000000	7.8848861	10.0000000	34.0000000

The resulting summary table clearly shows the [IQR](#) for the `points` variable, which is calculated as **13.00**. This value signifies that the middle half of the basketball players scored within a range of 13 points. Comparing this value to the [standard deviation](#) (8.22) and the overall range (34 - 10 = 24) provides a critical perspective on data variability. If the data were perfectly symmetrical and

normally distributed, the relationship between these measures would be tightly predictable; however, in this skewed performance dataset, the IQR offers a clearer insight into the typical spread.

## Going Deeper: Visualizing Quartiles (P25 and P75)

While the `QRANGE` option provides the final IQR value, it is often beneficial to examine the underlying components that define this range: the first quartile (Q1) and the third quartile (Q3). These quartiles represent the [25th percentile](#) and the [75th percentile](#), respectively. Knowing these exact cutoff points allows analysts to determine the precise boundaries of the central 50% of the data, which is crucial for building box plots or setting thresholds for [outlier](#) detection.

To include these specific quartile values in the output, [PROC MEANS](#) offers the dedicated keywords `P25` and `P75`. These options integrate seamlessly with the previously used statistic keywords, further enriching the summary table and providing a complete picture of the data's location and spread.

We enhance our [SAS](#) code one last time to include these [percentiles](#):

```
/*calculate summary statistics for points and include IQR*/
proc means data=my_data N Mean P25 P75 QRANGE Std Min Max;
var points;
run;
```

The final, fully detailed output is displayed below, featuring all requested [descriptive statistics](#). This table confirms the definition of the IQR: the `P25` (Q1) value is 12.00, and the `P75` (Q3) value is 25.00. Subtracting the former from the latter (25.00 - 12.00) yields the `QRANGE` value of 13.00, providing internal consistency and a complete statistical breakdown of the variable's distribution.

The MEANS Procedure							
Analysis Variable : points							
N	Mean	25th Pctl	75th Pctl	Quartile Range	Std Dev	Minimum	Maximum
15	17.8000000	12.0000000	25.0000000	13.0000000	7.8848861	10.0000000	34.0000000

## Conclusion: Enhancing Data Interpretation with IQR

Mastering the statistical options within [PROC MEANS](#) is an indispensable skill for any analyst working with [SAS](#). While default outputs offer a serviceable initial overview, customizing the

procedure to include robust measures of dispersion, such as the [IQR](#), significantly elevates the quality and reliability of your data summaries. The IQR provides a stable, non-parametric measure of spread that is particularly valuable when dealing with real-world data that deviates from the ideal normal distribution.

By consistently employing the `QRANGE` option, along with `P25` and `P75`, you move beyond relying solely on the mean and [standard deviation](#), which are sensitive to [outliers](#). This deeper, quartile-based understanding of data variability is critical for making informed decisions, accurately interpreting the central spread of your sample, and preparing data for subsequent advanced statistical modeling. We strongly encourage all SAS users to familiarize themselves with the full spectrum of options available in [PROC MEANS](#) to maximize their analytical efficiency and precision.

## Additional Resources for [SAS](#) Data Analysis

For those looking to expand their knowledge of advanced data summarization and preparation techniques within the SAS environment, the following tutorials provide guidance on other common tasks and essential procedures:

How to Calculate the [Coefficient of Variation](#) in SAS

How to Use [PROC FREQ](#) in SAS

How to Use [PROC SORT](#) in SAS