

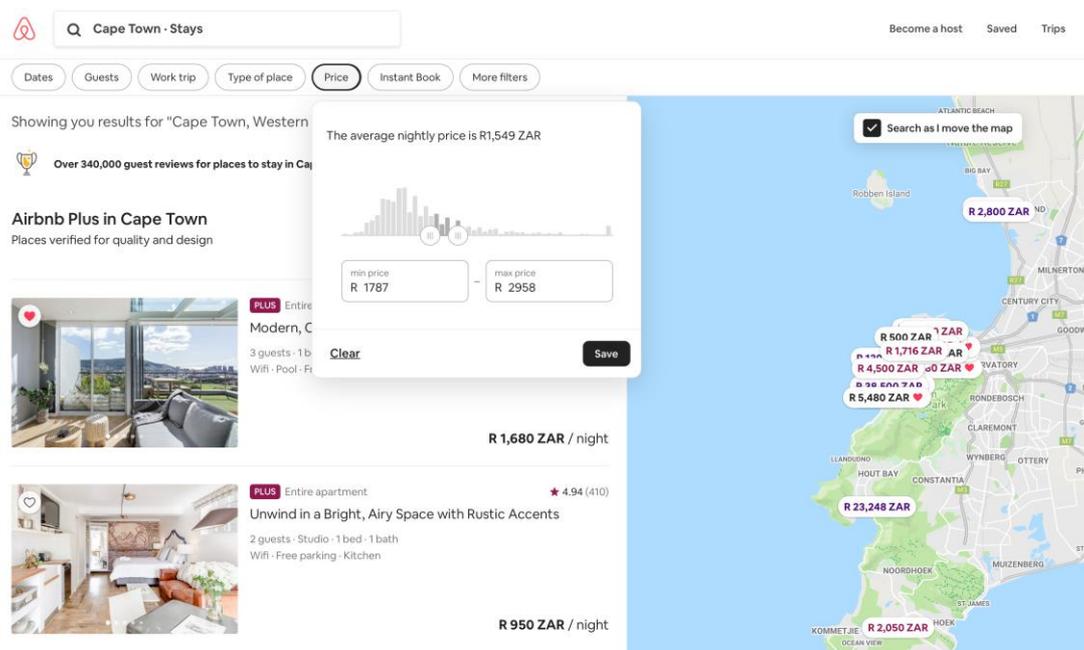
Statistics 101 for Reward Professionals

By Nicol Mullins, Deon Smit and Morag Phillips

Reward, numbers and statistics – three critical components for any reward professional

Reward professionals require a foundational understanding of statistics to analyse not only internal but also external data. It is the understanding of statistics that is key in the analysis, or when interpreting market data and comparing internal data to external information. Various statistics are essential to the reward professional. This paper provides an overview of the statistics required.

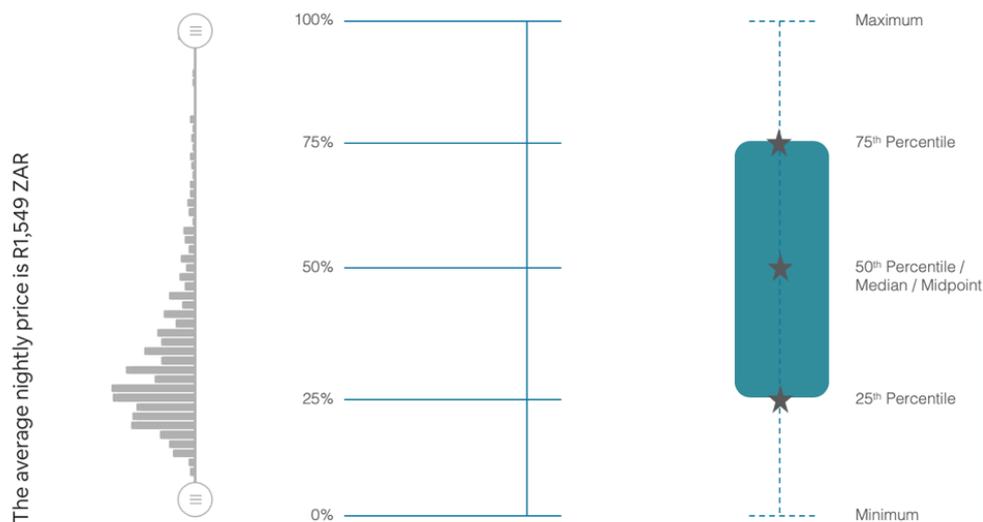
Some of the essential statistics will be minimum, maximum, market percentiles, which could be 25th, 50th (median/midpoint), 75th, or other percentiles, that are of value to the analysis. To illustrate the different statistical indicators that are of value we will use an Airbnb scale for accommodation in Cape Town. A range is provided on a slide representing the minimum, maximum values with the average being R1,549 per night.



Let's first review what a **percentile** is. It is a spread of the distribution of data. The most common definition of a percentile is a number where a certain percentage of scores fall below that percentile value. For example, the 25th percentile is the number where 25% of the observations are below this number and 75% are above. As an example, we could calculate percentiles of the height of people in a room, just the same as we could calculate the percentile spread of accommodation rates in a Cape Town suburb on Airbnb, or the spread of pay across Paterson B1 jobs in a mining company. It's just a calculation that paints a picture!

If we move on to **market salary percentiles** - market salary percentiles are an indicator of the distribution of what the market pays for specific roles or grades, when organisations participate in benchmark surveys, and the collected data is typically reported back on an aggregated basis. Percentiles typically represent this data. If one thinks of percentiles as

percentages, one can think of a range of zero per cent to 100% in compensation; no one can ever be paid zero per cent or a hundred per cent. It is, however, key to note that using this is only an illustration, as percentiles are not percentages and the two should not be confused nor used interchangeably. In the review of percentiles, we substitute these two percentages (0% and 100%), with **minimum** and **maximum**, and the values in between are considered the specific **range**, either for a grade, a job, a function or another metric in terms of evaluating a specific pay rate for a specific role or grade. These ranges would typically follow a bell curve distribution.



The 50th percentile, **median**, or midpoint, are synonyms for the same point in a dataset. The 50th percentile is the middle of the market set of data when ranked from high to low. When organisations benchmark on the 50th percentile, it means that they would be paying better than 50% of organisations and that 50% of organisations, would be paying better than that organisation in the sample. Subsequently, if you look at the **25th percentile**, it means that the organisation pay is better than 25% of companies in the data set. Benchmarking at the **75th percentile** would imply that our organisation pay is better than 75% of companies in the range. An easy way to understand the median is to visualise 101 people lined up in a row ranked with salaries from the lowest to the highest. The 50th percentile salary belongs to the person standing 50th in the row, where 50 people are earning less than the individual and 50 people are earning more than the individual. Remember that organisations don't pay everyone exactly the same salary, so where an organisation describes that they target the 50th percentile, it means that their own median (50th percentile) is close to the market 50th percentile.

A valuable additional statistic to also take into consideration when looking at market statistics is the average, also known as the **mean**. When your average is higher than the average for your market or the internal median, it would indicate that there are greater outliers at the top of the dataset which are pulling the average up. If the average is lower, there are outliers at the bottom. In other words, an average is skewed by outliers, whether high or low. For this reason, a median is a more robust reference point.

Reward professionals apply a mix of statistical calculations, dependent on what you are trying to reflect and on how much data you have.

	A	B	C	D	E
1					
2		Column	2	3	4
3			Calculation	Excel Formula	Comments
4		Employee ID	Basic Salary	Basic Salary	
5		Employee 1	100	100	
6		Employee 2	200	200	
7		Employee 3	400	300	
8					
9		Count	3	=COUNT(D7:D9)	<<< Number of observations in the data set
10		Min	100	=MIN(D7:D9)	<<< Minimum value in the data set
11		25th percentile	100	=PERCENTILE.EXC(D7:D9;0,25)	<<< 25th percentile in the data set
12		50th percentile	200	=PERCENTILE.EXC(D7:D9;0,5)	<<< 50th percentile in the data set
13		Average / Mean	233	=AVERAGE(D7:D9)	<<< Average value of the data set
14		75h percentile	400	=PERCENTILE.EXC(D7:D9;0,75)	<<< 75th percentile in the data set
15		Max	400	=MAX(D7:D9)	<<< Maximum value in the data set
16					
17					

If we were to report to Finance and apply the median, we would report that on grade 3 for employee's 1, 2 & 3 our median salary is R200. If Finance knows that we have 3 employees in grade 3, they might then multiply 200 with 3 to get a value of 600 for total costs. This would be incorrect as the sum of our salaries in grade 3 is 700—as such, using the average value for internal reporting is more useful than the median. Noting that the average value is higher than the median confirms that we have outliers to the top, which is observed in employee 3 being paid 400.

Remember that not all your readers will be comfortable with the full understanding of the statistics you are using. Therefore, an explanation is always helpful, as well as graphs. Have fun... numbers do indeed tell a thousand words!