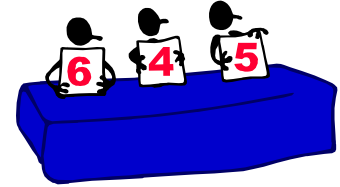


Case Study – HR Department Attribute Agreement Analysis

When to use Attribute Agreement Analysis

There are many business situations where people have to make a judgment about something.



Purchasing Process:	Is the amount on the invoice correct?
Finance Process:	Does this applicant meet all the criteria to qualify for a loan?
Manufacturing Process:	Is this part good or bad?
HR Process:	Do we classify this candidate as "Hire", "Possible" or "Decline".
Education:	Do all examiners award an essay the same grade?

We assume that experienced employees will make the right decision but how do we know that this is the case?

In a Six Sigma project, we need to collect process data in the search for root causes, but how do we know we have high quality data?

In both these situations, we need to use an Attribute Agreement Analysis to validate the capability of the decision making process.

Where the name comes from

You might classify a job applicant as **good** or **bad** – "good" and "bad" are attributes. Other examples....

The car is **green**

The unit is a **pass**

The unit is a **fail**

The amount is **incorrect**

The weather is **hot**

In an Attribute Agreement Analysis we look at the attribute assigned to a particular item by different people, by the same person on different occasions and the "right answer" and determine the level of agreement. Obviously we would like to have a high level of agreement but it's surprising how often we don't!

Case Study – HR Department Attribute Agreement Analysis

How to do it

You need to set up a structured study where a number of items will be assessed more than one time by more than one assessor.

This case study shows how a newly trained Black Belt in an HR Department used the technique to ensure interviewers were correctly assessing candidates for technical support positions.

There were a large number of applicants and a team of interviewers was going to be used to carry out initial screening interviews by telephone. The interviewers' task was to categorise the applicants as Pass or Fail (the attribute was Pass or Fail). Before proceeding, an Attribute Agreement Analysis was carried out.

The study could have two outcomes

1. the interviewers always (or almost always) get it right
2. they get it wrong too often

In the first case you would be able to proceed with confidence – just imagine what that would feel like - no nagging doubts about whether the interviewers had done a good job but confidence that the best applicants really were being selected.

In the second case you would have a problem and there would be no point in going ahead with the interviews until you resolved it. Fortunately, the output of the Attribute Agreement Analysis gives pointers as to where the interviewers are getting it wrong and this information leads to taking appropriate actions to improve their capability.

Setting up and running the study

The study should always reflect real life as closely as possible. Here, a number of interviews were recorded and the interviewers were asked to listen to these and decide if the candidate was a Pass or a Fail. The procedure was as follows.

Case Study – HR Department Attribute Agreement Analysis

1. Thirty recordings were selected. Roughly half were Pass and half Fail. These included some obvious Passes, some obvious Fails and the rest were spread in-between including some borderline cases.
2. A panel of experts verified the classification of each of the candidates. This gave the “right answer”.
3. Three interviewers were selected to take part in the study. Each interviewer listened to each recording in turn and assigned the candidate Pass or Fail. The interviewers weren’t allowed to know what attribute the other interviewers gave – to avoid bias (influencing each other).
4. Now Step 3 was repeated but we didn’t want the interviewers to remember how they classified each candidate last time. To help ensure this, the calls were randomly reordered. Also, the second assessment was held a week later so there was little chance of the interviewers remembering which candidate was which or how they rated each candidate initially.
5. The study was now complete and the data could be analysed.

Results Table

Candidate	Right	Jan1	Jan2	Chris1	Chris2	Sam1	Sam2
1	F	F	F	F	F	F	F
2	P	P	P	P	P	F	F
3	F	F	F	F	F	F	F
4	P	P	P	P	P	P	P
5	P	P	P	P	P	F	P
6	F	P	P	P	F	F	F
7	F	F	F	F	F	F	F
8	F	P	P	F	F	F	F
9	P	P	P	P	P	F	F
10	P	P	P	P	P	F	F
11	F	F	F	F	F	F	F
12	F	P	P	F	F	F	F
13	F	F	F	F	F	F	F
14	F	P	P	F	F	F	F
15	P	P	P	P	F	P	P
16	P	P	P	P	P	P	P
17	P	P	P	P	P	P	P
18	P	P	P	P	P	P	P
19	P	P	P	P	P	P	P
20	P	P	P	P	P	F	F
21	F	P	P	F	F	F	F
22	F	P	P	F	F	F	F
23	P	P	P	P	P	P	P

Case Study – HR Department Attribute Agreement Analysis

24	F	P	F	F	F	F	F
25	F	F	P	F	F	F	F
26	F	F	F	F	F	F	F
27	F	F	F	F	F	F	F
28	P	P	P	P	P	F	F
29	P	P	P	P	P	P	P
30	F	F	F	F	F	F	F

Key to Column Names

Candidate = Candidate Number (arbitrary)

Right = The expert assessment of that candidate

Jan1 = Jan's assessment the first time she assessed each candidate

Jan2 = Jan's assessment the second time she assessed each candidate

Chris1, Chris2 = Chris's first and second assessment for each candidate

Sam1, Sam2 = Sam's first and second assessment for each candidate

P = Pass, F = Fail

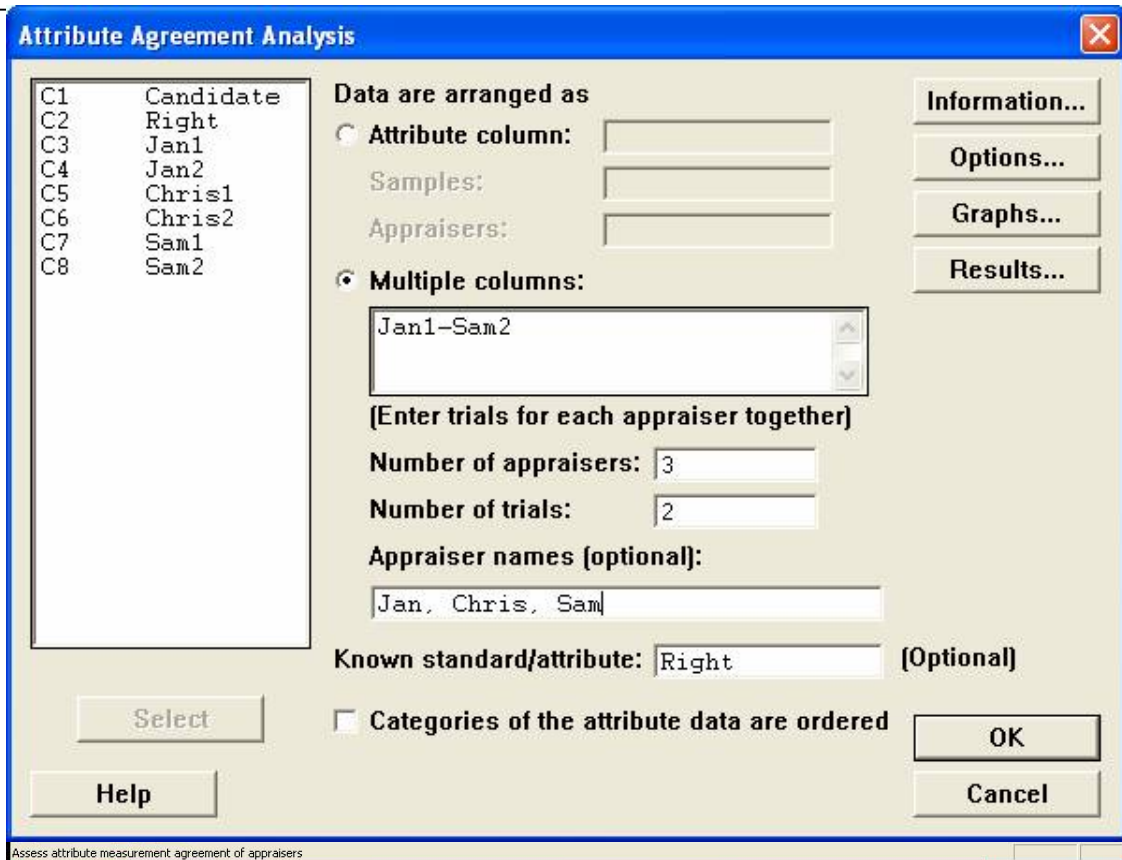
Note that the order of data in the worksheet is not the order in which the study was conducted (it was run in random order as described above).

Analysing the study

The data is best analysed using statistical software. Here we used Minitab (if you would like an Excel worksheet that works for simple cases please get in touch).

Stat>Quality Tools>Attribute Agreement Analysis

Case Study – HR Department Attribute Agreement Analysis



Attribute Agreement Analysis

C1	Candidate
C2	Right
C3	Jan1
C4	Jan2
C5	Chris1
C6	Chris2
C7	Sam1
C8	Sam2

Data are arranged as

Attribute column: []

Samples: []

Appraisers: []

Multiple columns:

[Jan1-Sam2]

[Enter trials for each appraiser together]

Number of appraisers: [3]

Number of trials: [2]

Appraiser names (optional): [Jan, Chris, Sam]

Known standard/attribute: [Right] (Optional)

Categories of the attribute data are ordered

Buttons: Select, Help, Information..., Options..., Graphs..., Results..., OK, Cancel

Assess attribute measurement agreement of appraisers

The dialogue box below opens

Our data is in multiple columns so select this arrangement.

Drag the columns containing the data into the data field.

Tell Minitab that there were 3 appraisers (assessors) and 2 trials (each assessor classified each call twice).

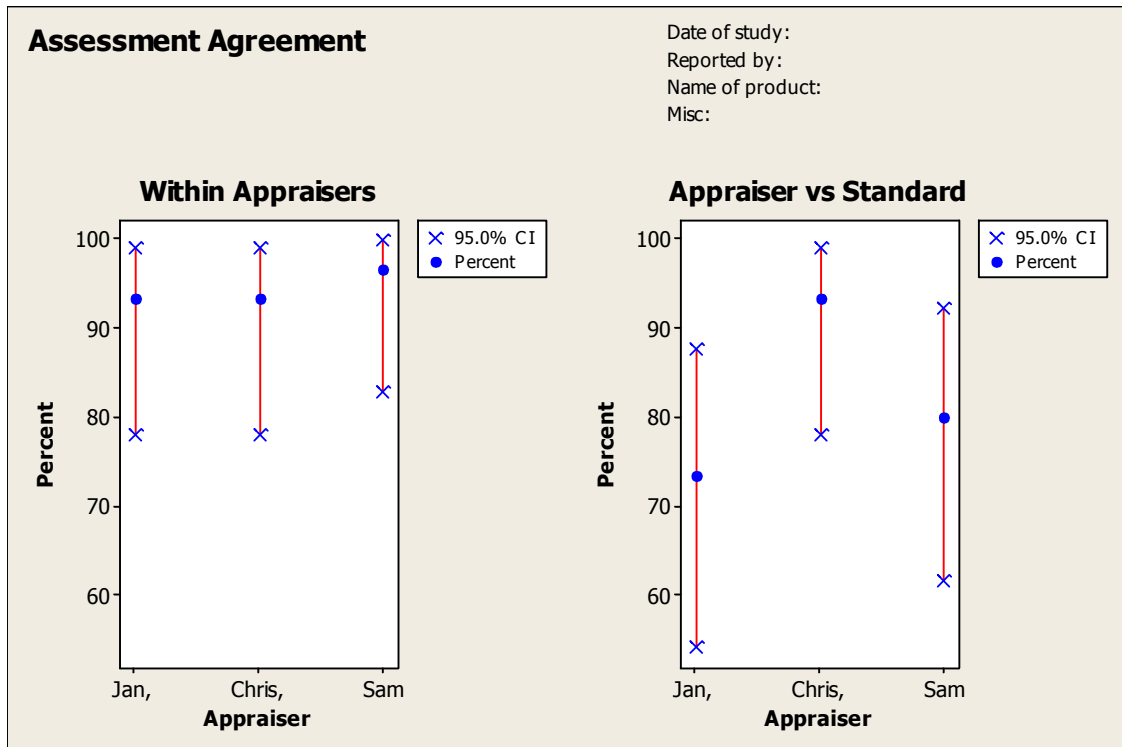
Note that Minitab assumes that the first 2 column names are the first and second assessment by Interviewer 1, third and fourth are Interviewer 2 etc.

Optionally, input the appraiser names

Also tell Minitab that the known standard is contained in the column Right

Hit OK and you get the graph shown below.

Case Study – HR Department Attribute Agreement Analysis



Left graph

There are 3 vertical lines, one for each appraiser. The blue dot shows how well the appraiser agreed with themselves across the two assessments made on each candidate. Jan's two assessments agreed with each other about 93% of the time. If you look at the table of results, you will see that on 2 occasions she disagreed with her own previous assessment – on Candidate 24 and Candidate 25. The red line which extends above and below the blue dot shows the 95% confidence interval. Confidence intervals can be large when working with attribute data but they can be reduced by using higher sample sizes – in this case that would mean assessing more than 30 candidates.

In general, this graph shows that all three appraisers were quite consistent with their own previous judgment.

Right graph

Case Study – HR Department Attribute Agreement Analysis

The dots this time show how well each appraiser agreed with the right answer. Around 73% of the time, both of Jan’s assessments of the candidates agreed with the right answer. We already know from the left graph that Jan agreed with herself most of the time so the conclusion is that she disagreed with the right answer quite frequently.

In general, this graph shows that two of the interviewers disagreed frequently with the right answer. There was a problem!

If we now look at Minitab’s session window output we can get some more detailed information. We’ll look at it bit by bit...

Within Appraisers

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95 % CI
Jan,	30	28	93.33	(77.93, 99.18)
Chris,	30	28	93.33	(77.93, 99.18)
Sam	30	29	96.67	(82.78, 99.92)

Matched: Appraiser agrees with him/herself across trials.

The first section (above) is a numerical version of what we saw in the left graph.

Each Appraiser vs Standard

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95 % CI
Jan,	30	22	73.33	(54.11, 87.72)
Chris,	30	28	93.33	(77.93, 99.18)
Sam	30	24	80.00	(61.43, 92.29)

Matched: Appraiser's assessment across trials agrees with the known standard.

Assessment Disagreement

Appraiser	# P / F	Percent	# F / P	Percent	# Mixed	Percent
Jan,	6	37.50	0	0.00	2	6.67
Chris,	0	0.00	0	0.00	2	6.67
Sam	0	0.00	5	35.71	1	3.33

P / F: Assessments across trials = P / standard = F.

F / P: Assessments across trials = F / standard = P.

Case Study – HR Department Attribute Agreement Analysis

Mixed: Assessments across trials are not identical.

The top half of the next section (above) is a numerical version of the right graph. The bottom half (Assessment Disagreement) adds some useful information. On 6 occasions, Jan classified a failed candidate as a pass, but on no occasions did she classify a Pass as a Fail. In other words, in general, she was “too lenient”; she would tend to pass poor candidates. The mixed column shows 2 other occasions when she was inconsistent in her judgment – we already know about these (Candidates 24 and 25).

Chris did not consistently mis-classify any candidates but Sam classified a Pass as a Fail on 5 occasions – in other words, he was “too tough”, he would sometimes fail good candidates.

This has given us great insight into the problem but let’s look at the rest of Minitab’s information before we draw our final conclusions.

Between Appraisers

Assessment Agreement

# Inspected	# Matched	Percent	95 % CI
30	15	50.00	(31.30, 68.70)

Matched: All appraisers' assessments agree with each other.

This section (above) looks at how well the 3 interviewers agree with each other. We can see that for 15 candidates, all 3 agreed with each other on both assessments (irrespective of being right or wrong). Thinking about what we already know, this is not that surprising – we have already seen that Jan and Sam had different opinions on what makes a good or bad candidate.

All Appraisers vs Standard

Assessment Agreement

# Inspected	# Matched	Percent	95 % CI
30	15	50.00	(31.30, 68.70)

Matched: All appraisers' assessments agree with the known standard.

The last section of Minitab output (above) tells us that for 15 out of 30 candidates, all the interviewers gave the correct answer both times

Case Study – HR Department Attribute Agreement Analysis

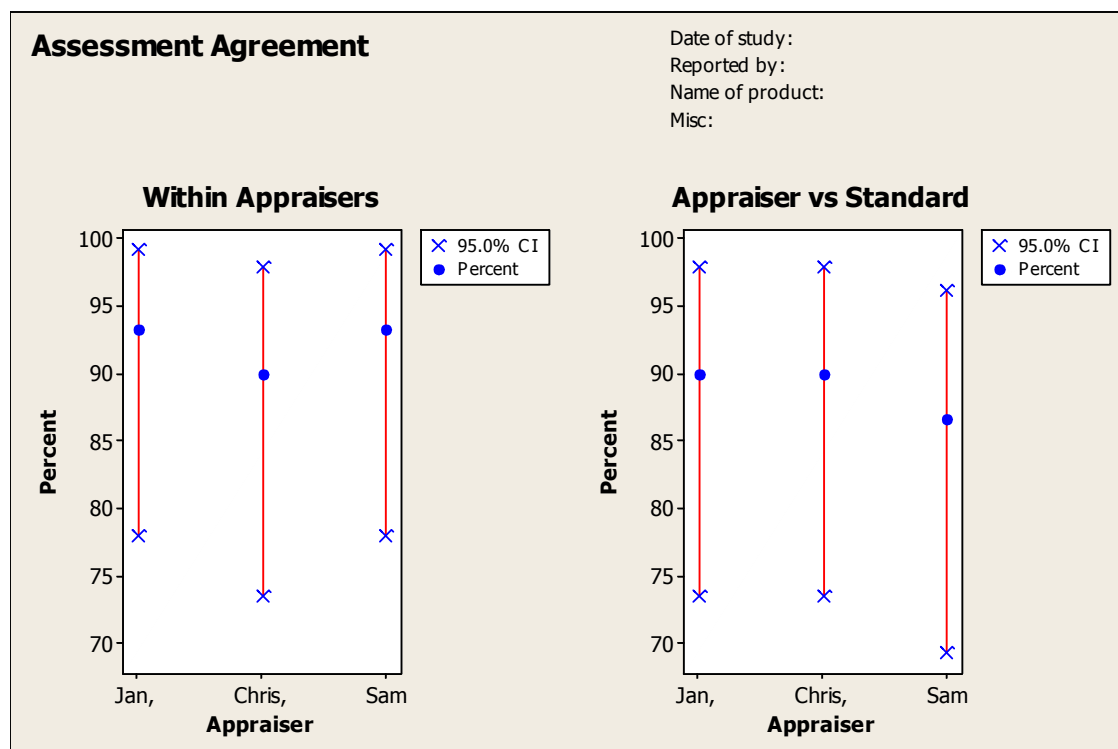
they assessed them. (In this case the values are same as Between Appraisers but that will not always be so).

Conclusion of this study

The results showed that there was a problem with the ability of the interviewers to assess candidates. Having ascertained this, the study gave us evidence of where the problem originated. There was not a serious issue with interviewers being inconsistent when assessing the same candidate, but the three interviewers were applying the selection criteria differently. This was probably a training issue.

It didn't take too much investigation to ascertain that, in fact, no training had been given – after all these were experienced interviewers, this was their job!

Remember that these three interviewers represented all the interviewers so there were likely to be others with similar problems. Everyone was retrained and then the study was repeated to confirm capability.



Case Study – HR Department Attribute Agreement Analysis



The results showed a marked improvement and now the interviews were able to proceed.

We have to accept that perfection is impossible and there will always be some level of disagreement, but while you should continuously try to improve, you also need to recognise when you have reached good capability (like we did here) and can proceed with the task in hand with high confidence in the results.

Final note: Although in this case training resolved the problem, training is not always the solution. In fact, the most common problem with attribute assessment is lack of, or ambiguous, decision criteria. If this is a problem, a study like this one will reveal that, too.

Final Comments

Every manager has to make decisions. To make informed decisions requires information but it has to be good information. Attribute Agreement Analysis is a simple technique to validate the quality of process information and to provide direction to finding problems where they exist.