DATA ANALYSIS

Developing Professional Practice and Using Information in HR

Module Code: MOD001181
Session outline

- Why don’t Business leaders see HR as integral to strategy?
- What data does HR collate?
- What can be done with this data?
- Identify and scrutinize appropriate HR data sources
HR role

Previously the HR function played a role that was

- Largely administrative
- Carried a range of basic tasks – issuing contracts, maintaining records, payroll etc.

HRM is about – creating an efficient infrastructure and also mobilising an effective set of differentiators that together make an impressive difference to the quality of the organization and its resulting performance. (Taylor and Woodhams 2012, page 77)
Good HR

Good HR is strategic and will focus on ‘adding value’ through:

1. Maintenance activities
2. Crisis prevention activities
3. Continuous improvement activities
4. Change management activities.
5. And being ‘business savvy’
   - Understanding business model in depth
   - Generating insight through evidence and data
   - Connecting with curiosity, purpose and impact
   - Leading with integrity, consideration and challenge.
Why is HR not generally seen as integral to strategy:

- CIPD survey shows two in every three HR leaders believe they play an integral role in strategy making at the organisations where they work. The bad news is that only one in three business leaders agrees with them.
- A historical legacy and HR needs to be more data-savvy and take notice of the three Vs – volume, variety and velocity.
- There’s no point in collecting the data if you don’t use it and use it effectively.
- Board – analytics, HR data more focused on qualitative
  - Five C’s – Cost, Compliance, Capacity, Connections, Capability
Terry Leahy when he was chief executive of Tesco said,

“When the marketing manager comes to me he brings insightful and predictive information about our customers. We know what our customers are thinking and feeling, what they’re going to do, how they see Tesco in relation to their needs and wants. It’s hugely insightful we make decisions on it.” So that's what he said about marketing. … “When the HR manager comes to our meetings it’s not predictive, it’s not insightful, it’s rear view mirror and we can't make decisions on it.”

You need to start to take the information that the HR function has and I think to really drive insight you need to explore it, you need to work with it, you need to integrate it with other sources of information and it’s only then will it start to tell a story.
High commitment for HR

Arguments

- High Commitment HR practices lead to:
  - Improved attitudes and behaviours
  - Lower absenteeism and labour turnover
  - High levels of productivity, quality and customer service

Resulting in Improved Organisational Performance
Data Collection

- Compilation and interpretation of primary and secondary sources of information.

- The integration of different sources will consolidate the write up of the report.
Primary Source
• Data is collected by researcher himself
  • Data is gathered through questionnaire, interviews, observations etc.

Secondary Source
• Data collected, compiled or written by other researchers eg. books, journals, newspapers
  • Any reference must be acknowledged
STEPS TO COLLECT DATA

**REVIEW & COMPILE SECONDARY SOURCE INFORMATION**
(Referred to in the BACKGROUND/INTRODUCTION section of report)

**PLAN & DESIGN DATA COLLECTION INSTRUMENTS TO GATHER PRIMARY INFORMATION**
(Referred to in the FINDINGS, CONCLUSIONS & RECOMMENDATIONS sections of report)

**DATA COLLECTION**

**DATA ANALYSIS AND INTERPRETATION**
Effective way of gathering information

Involves verbal and non-verbal communications

INTERVIEW

Can be conducted face to face, by telephone, online or through mail
Steps To An Effective Interview

Prepare your interview schedule

Select your subjects/ key informants

Conduct the interview

Analyze and interpret data collected from the interview
Drawback to interviews

- They can consume a great deal of time if interviewers take full advantage of the opportunity to hear respondents out and change their questions accordingly.
- Personal biases can also distort the data.
- The nature of the question and the interactions between the interviewer and the respondent may discourage or encourage certain kinds of responses.
- It take considerable skill to gather valid data.
The most common data collection instrument

Useful to collect quantitative and qualitative information

Survey Questionnaire

Should contain 3 elements:
1. Introduction – to explain the objectives
2. Instructions – must be clear, simple language & short
3. User-friendly – avoid difficult or ambiguous questions
Steps To An Effective Survey Questionnaire

**Prepare your survey questions**
(Formulate & choose types of questions, order them, write instructions, make copies)

**Select your respondents/sampling**
Random/Selected

**Administer the survey questionnaire**
(date, venue, time)

**Tabulate data collected**
(Statistical analysis-frequency/mean/correlation/%)

**Analyze and interpret data collected**
Questionnaires; there are drawbacks;

- Responses are limited to the questions asked in the instrument.
- They provide little opportunity to probe for additional data or ask for points of clarification.
- They tend to be impersonal.
- Often elicit response biases – tend to answer in a socially acceptable manner.
Observe verbal & non-verbal communication, surrounding atmosphere, culture & situation

Need to keep meticulous records of the observations

Observations

Can be done through discussions, observations of habits, rituals, review of documentation, experiments
Steps To An Effective Observation

Determine what needs to be observed
(Plan, prepare checklist, how to record data)

Select your participants
Random/Selected

Conduct the observation
(venue, duration, recording materials, take photographs)

Compile data collected

Analyze and interpret data collected
Advantages to Observation:

- They are free of the biases inherent in the self-report data.
- They put the practitioner directly in touch with the behaviors in question.
- They involved real-time data, describing behavior occurring in the present rather than the past.
- They are adapting in that they can be modified depending on what is being observed.
Problems with Observation

- Difficulties interpreting the meaning underlying the observations.
- Observers must decide which people to observe; choose time periods, territory and events
- Failure to attend to these sampling issues can result in a biased sample of data.
3. In a small scale study, the most common forms of statistical analysis presented are:
   • Frequency
   • Mean
   • Percentage
DATA INTERPRETATION

1. It involves 2 terms
   • ‘Results’ – presentation of data/findings (statistics)
   • ‘Discussion’ – interpretation of data/findings

2. Things to consider when interpreting your data:
   • Interpret findings based on the purpose and objectives of your study
   • Relate the findings to real life context
   • Use persuasive language to convince your readers to see the research from your point of view.
   • Order your interpretation to highlight the most important findings
   • Include limitations to your research.
   • Use simple, clear language
CODING CLOSED QUESTIONS

- You will need to consider ways to code your data for entry into a computer system. Each response to a question needs an individual code that allows for data analysis. For example, a common question is ‘What gender are you?’ This could be coded as Male or Female in a spreadsheet. However, subsequent analysis is easier and more accurate if a code is used – say:
  1   Male
  2   Female.
- There must also be a code for a non-response or a missing response, and with gender it is sometimes useful to know the reason for a non-response. So the full coding of the question would be:
  1   Male
  2   Female
  3   Would prefer not to respond
  4   Is unwilling to be characterised in this manner
  5   This information is confidential
  9   Missing

Note that 9, 99, 999 or 9999 is ordinarily used as the code for missing data.
CODING CLOSED QUESTIONS

• If you are intending to use a secondary data source and carry out comparative analysis with that source, then coding your own research using the same coding system will make the comparison considerably easier.

• If you are collecting data on the socio-economic class of your respondents, you could use the standard national statistics classification. Then any subsequent analysis with other data sets will be easier.

**Analytic socio-economic classes**

1 Higher managerial and professional occupations
   1.1 Large employers and higher managerial occupations
   1.2 Higher professional occupations

2 Lower managerial and professional occupations

3 Intermediate occupations

4 Small employers and own-account workers

5 Lower supervisory and technical occupations

6 Semi-routine occupations

7 Routine occupations

8 Never worked, and long-term unemployed.
CODING CLOSED QUESTIONS

• For a scaled question such as ‘How enjoyable did you find your university studies? Answer on the scale 1–6’, answer codes might be:
  Very enjoyable 1
  Quite enjoyable 2
  Mostly enjoyable 3
  Not very enjoyable 4
  Not enjoyable 5
  Dreadful 6

• **Likert scales** are used to test the agreement or disagreement of a respondent to a proposition. A Likert scale question is coded in the same way as other scaled questions.
  Agree 1
  Tend to agree 2
  Tend to disagree 3
  Disagree 4
CODING CLOSED QUESTIONS

- If you are using survey or questionnaire software, such as Snap, the data coding is done automatically.

- When the data is entered into a spreadsheet the cells will contain only the code number. A narrative reminder of the coding of the question is often present in the column heading.

- When using SPSS you set up the variables and their characteristics as you enter the data.
What sort of things might be entered in Other? Such things as:

- My partner moved jobs to a different part of the country.
- The transport links to my old job were just too poor.
- I wanted a shorter journey to work.
- Two of my friends worked at the same company.
- My mum is the managing director.
- I worked here before and I really liked it.
- This job is completely different from my last job. I was a teacher and now I am an administrator.

The most common way to deal with open data is to create code categories for the responses, so that the data above is entered into electronic form as:
Geographical change of area 11
Transport problems 12
Shorter journeys to work 13
Social connection to existing workers 14
Kinship relationship to existing worker 15
Previously employed at the organisation and liked it 16
Change of career 17
Validating data: FIELD EDITING

• There are two stages to validating data: field editing and office editing.

• Field editing refers to monitoring the actual collection of data ‘in the field’. In dissertation research you are very likely to have to collect your data by means of face-to-face personal contact. Field editing, or checking the validity of the data at the point of entry, is vital.

• The key issues in field editing are:
  • to get accurate answers
  • to record an answer to every question
  • to ensure that any writing is legible.

• If you are asking the respondent the questions, plan to keep about two to three minutes between respondents so that you can check and correct the entries on the completed form.

• You will be surprised how often after only a few seconds you cannot quite read or remember the answer just given.
Validating data: FIELD EDITING

• When the respondent is self-completing the questionnaire, you must check the form before the person walks away.
• It should only take a minute to look over the form but if anything is not clear or is missing, you will have the chance to ask if your respondent missed the question or chose not to respond.
• You may also need to ask the question: ‘What does that say?’
• Check through this list after every respondent’s questionnaire completion:
  • All questions have been answered or a non-response given.
  • Answers are in the correct format.
  • Inappropriate answers have been identified and corrected
  • Answers in the wrong place on the form have been replaced correctly.
  • Open questions’ answers are legible.
  • Biographical data is present and ‘visually’ looks correct.
  • Any note the interviewer has made is legible.
Validating data: OFFICE EDITING

- Office editing involves a more thorough check of the questionnaire for the same items as in the field editing checklist:
  - All questions have been answered or a non-response given
  - Answers are in the correct format.
  - Inappropriate answers have been identified and corrected.
  - Answers in the wrong place on the form have been replaced correctly
  - Open questions’ answers are legible.
  - Biographical data is present and ‘visually’ looks correct.
  - Any note the interviewer has made is legible.

- This editing process can make small corrections to ‘obvious’ errors such as marking the answers in the wrong format, or putting the answer in the wrong place, or using the wrong data format.

- But this editing process should never include adding missing answers that you think the respondent would have made – in academic dissertations this would be regarded as a form of plagiarism.
Validating data: OFFICE EDITING

• Errors can occur when data is entered and once entered are quite difficult to find and remove. Standard transposition errors can occur frequently and create havoc in your data.

• A transposition error occurs where the entry of two digits is reversed.

• Because questionnaires use codes for data, there are plenty of opportunities to create such havoc. The code 12 is certainly a possible code for a question with a large number of possible answers. The code 21 is not likely to represent useful data. When entering data on a numeric keypad it is very easy to enter 21 for the code 12.

• Most of the data analysis software can set up input masks that significantly reduce the possibility of entering data wrongly.

• An input mask restricts the data that can be entered to a predetermined range. So if a question has 14 answers and uses 15 for missing data, the input mask can restrict the entry of data in that cell to the range 1–15.
Questionnaire data analysis is normally undertaken using one of the following software products:

- Excel spreadsheet
- Access database
- survey software, such as Snap
- SPSS software
Example of an Excel spreadsheet

<table>
<thead>
<tr>
<th>ID No</th>
<th>Family Name</th>
<th>First Name</th>
<th>Sample</th>
<th>Gender</th>
<th>Age</th>
<th>Q1 No you like your job, 1-5</th>
<th>Q2 How long have you been in this job, years</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Smith</td>
<td>Fred</td>
<td>TRUE</td>
<td>1</td>
<td>67</td>
<td>1</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Frost</td>
<td>Pat</td>
<td>TRUE</td>
<td>1</td>
<td>23</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Buckley</td>
<td>Susan</td>
<td>TRUE</td>
<td>2</td>
<td>19</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Nichols</td>
<td>Dandy</td>
<td>TRUE</td>
<td>2</td>
<td>42</td>
<td>1</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Farmer</td>
<td>Peter</td>
<td>TRUE</td>
<td>1</td>
<td>28</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example of a Snap software data entry

![Image: Snap software interface showing data entry for a book buying survey.](image)
Example of a Snap variable view sheet