In this project, we'll play the role of data analyst and pretend our stakeholders want to know the following: 1- Are employees who only worked for the institutes for a short period of time resigning due to some kind of dissatisfaction? 2- What about employees who have been there longer? 3- Are younger employees resigning due to some kind of dissatisfaction? What about older employees? They want us to combine the results for both surveys to answer these questions. However, although both used the same survey template, one of them customized some of the answers. In the guided steps, we'll aim to do most of the data cleaning and get you started analyzing the first question.

In [1]:
```
import pandas as pd
import numpy as np
# Read the tafe_survey.csv CSV file into pandas, and assign it to the v
ariable name dete_survey.
tafe_survey = pd.read_csv('tafe-employee-exit-survey-access-database-de
cember-2013.csv', encoding= "cp1252")

#Quick exploration of the data

pd.options.display.max_columns = 150 # to avoid truncated output
tafe_survey.head()
```

Out[1]:

	Record ID	Institute	WorkArea	CESSATION YEAR	Reason for ceasing employment	Contributing Factors. Career Move - Public Sector	Contribut Facto Car Mov Priv Sec
0	634133009996094000	Southern Queensland Institute of TAFE	Non-Delivery (corporate)	2010.0	Contract Expired	NaN	N

1	634133654064531000	Mount Isa Institute of TAFE	Non-Delivery (corporate)	2010.0	Retirement	-	
2	634138845606563000	Mount Isa Institute of TAFE	Delivery (teaching)	2010.0	Retirement	-	
3	634139903350000000	Mount Isa Institute of TAFE	Non-Delivery (corporate)	2010.0	Resignation	-	
4	634146578511788000	Southern Queensland Institute of TAFE	Delivery (teaching)	2010.0	Resignation	-	Career M - Priv Se

In [2]:
```
tafe_survey.info
```

Out[2]:
```
<bound method DataFrame.info of              Record ID
                Institute  \
0    634133009996094000  Southern Queensland Institute of TAFE
1    634133654064531000            Mount Isa Institute of TAFE
2    634138845606563000            Mount Isa Institute of TAFE
3    634139903350000000            Mount Isa Institute of TAFE
4    634146578511788000  Southern Queensland Institute of TAFE
..                  ...                                    ...
697  635066785175197000         Barrier Reef Institute of TAFE
698  635067716405666000  Southern Queensland Institute of TAFE
699  635070442972541000       Tropical North Institute of TAFE
700  635071153170979000      Southbank Institute of Technology
701  635073030973791000       Tropical North Institute of TAFE

                     WorkArea  CESSATION YEAR Reason for ceasing employ
ment  \
0    Non-Delivery (corporate)          2010.0              Contract Exp
ired
1    Non-Delivery (corporate)          2010.0                   Retire
ment
2         Delivery (teaching)          2010.0                   Retire
ment
```

```
3    Non-Delivery (corporate)          2010.0                    Resigna
tion
4         Delivery (teaching)          2010.0                    Resigna
tion
..                        ...             ...                         
...
697       Delivery (teaching)          2013.0                    Resigna
tion
698  Non-Delivery (corporate)          2013.0                    Resigna
tion
699       Delivery (teaching)          2013.0                    Resigna
tion
700  Non-Delivery (corporate)          2013.0              Contract Exp
ired
701  Non-Delivery (corporate)          2013.0                    Resigna
tion

    Contributing Factors. Career Move - Public Sector   \
0                                                  NaN
1                                                    -
2                                                    -
3                                                    -
4                                                    -
..                                                 ...
697                       Career Move - Public Sector
698                       Career Move - Public Sector
699                                                  -
700                                                NaN
701                                                  -

    Contributing Factors. Career Move - Private Sector   \
0                                                  NaN
1                                                    -
2                                                    -
3                                                    -
4                         Career Move - Private Sector
..                                                 ...
697                                                  -
698                                                  -
```

```
699                                                    -
700                                                  NaN
701                                                    -

    Contributing Factors. Career Move - Self-employment  \
0                                                    NaN
1                                                      -
2                                                      -
3                                                      -
4                                                      -
..                                                   ...
697                                                    -
698                                                    -
699                                                    -
700                                                  NaN
701                       Career Move - Self-employment

    Contributing Factors. Ill Health Contributing Factors. Maternity/Fa
mily  \
0                                  NaN
NaN
1                                    -
   -
2                                    -
   -
3                                    -
   -
4                                    -
   -
..                                 ...
...
697                                  -
   -
698                                  -
   -
699                                  -
   -
700                                NaN
NaN
```

```
701                                  -
   -

    Contributing Factors. Dissatisfaction  \
0                                       NaN
1                                         -
2                                         -
3                                         -
4                                         -
..                                      ...
697                                       -
698                                       -
699                                       -
700                                     NaN
701                                       -

    Contributing Factors. Job Dissatisfaction  \
0                                         NaN
1                                           -
2                                           -
3                                           -
4                                           -
..                                        ...
697                                         -
698                                         -
699                                         -
700                                       NaN
701                                         -

    Contributing Factors. Interpersonal Conflict Contributing Factors.
Study  \
0                                             NaN
  NaN
1                                               -
    -
2                                               -
    -
3                                               -
    -
```

```
4                                                   -
    -
..                                                ...
  ...
697                                                 -
    -
698                                                 -
    -
699                                                 -
    -
700                                               NaN
  NaN
701                                                 -
    -

    Contributing Factors. Travel Contributing Factors. Other  \
0                           NaN                           NaN
1                        Travel                             -
2                             -                             -
3                        Travel                             -
4                             -                             -
..                          ...                           ...
697                           -                             -
698                           -                             -
699                           -                         Other
700                         NaN                           NaN
701                      Travel                             -

    Contributing Factors. NONE  \
0                          NaN
1                            -
2                         NONE
3                            -
4                            -
..                         ...
697                          -
698                          -
699                          -
700                        NaN
```

```
701                          -

    Main Factor.     Which of these was the main factor for leaving?  \
0                                                            NaN
1                                                            NaN
2                                                            NaN
3                                                            NaN
4                                                            NaN
..                                                           ...
697                                                          NaN
698                                                          NaN
699                                                          NaN
700                                                          NaN
701                        Career Move - Self-employment

    InstituteViews. Topic:1. I feel the senior leadership had a clear v
ision and direction  \
0                                                    Agree

1                                                    Agree

2                                                    Agree

3                                                    Agree

4                                                    Agree

..                                                     ...

697                                                Neutral

698                                                  NaN

699                                                  Agree

700                                       Strongly Agree

701                                       Strongly Agree
```

```
    InstituteViews. Topic:2. I was given access to skills training to h
elp me do my job better  \
0                                                    Agree

1                                                    Agree

2                                                    Agree

3                                                    Agree

4                                                    Agree

..                                                     ...

697                                                  Agree

698                                                    NaN

699                                       Strongly Agree

700                                    Strongly Disagree

701                                       Strongly Agree


    InstituteViews. Topic:3. I was given adequate opportunities for per
sonal development  \
0                                                    Agree

1                                                    Agree

2                                                    Agree

3                                                    Agree

4                                       Strongly Agree

..                                                     ...
```

```
697                                                 Agree

698                                                   NaN

699                                            Strongly Agree

700                                         Strongly Disagree

701                                            Strongly Agree


    InstituteViews. Topic:4. I was given adequate opportunities for pro
motion within %Institute]Q25LBL%  \
0                                                   Neutral

1                                                     Agree

2                                                     Agree

3                                                     Agree

4                                                     Agree

..                                                      ...

697                                                 Neutral

698                                                   NaN

699                                                 Neutral

700                                         Strongly Disagree

701                                            Strongly Agree


    InstituteViews. Topic:5. I felt the salary for the job was right fo
r the responsibilities I had  \
```

```
0                                                  Agree

1                                                  Agree

2                                                  Agree

3                                                  Agree

4                                           Strongly Agree

..                                                    ...

697                                            Disagree

698                                                 NaN

699                                                Agree

700                                            Disagree

701                                         Strongly Agree


    InstituteViews. Topic:6. The organisation recognised when staff did
good work  \
0                                                  Agree

1                                           Strongly Agree

2                                                  Agree

3                                                  Agree

4                                           Strongly Agree

..                                                    ...

697                                             Neutral
```

```
698                                                  NaN

699                                                Agree

700                                                Agree

701                                     Strongly Agree


    InstituteViews. Topic:7. Management was generally supportive of me
\
0                                                  Agree

1                                       Strongly Agree

2                                       Strongly Agree

3                                                  Agree

4                                       Strongly Agree

..                                                   ...

697                                                Agree

698                                                  NaN

699                                                Agree

700                                              Neutral

701                                     Strongly Agree


    InstituteViews. Topic:8. Management was generally supportive of my
team  \
0                                                  Agree

1                                                  Agree
```

```
2                                                    Agree

3                                                    Agree

4                                           Strongly Agree

..                                                     ...

697                                                  Agree

698                                                    NaN

699                                                  Agree

700                                                Neutral

701                                         Strongly Agree


    InstituteViews. Topic:9. I was kept informed of the changes in the
organisation which would affect me  \
0                                                    Agree

1                                           Strongly Agree

2                                                    Agree

3                                                    Agree

4                                                    Agree

..                                                     ...

697                                                  Agree

698                                                    NaN

699                                                  Agree
```

```
700                                          Strongly Agree

701                                          Strongly Agree


    InstituteViews. Topic:10. Staff morale was positive within the Inst
itute  \
0                                                    Agree

1                                                    Agree

2                                                    Agree

3                                                    Agree

4                                            Strongly Agree

..                                                       ...

697                                             Disagree

698                                                  NaN

699                                              Neutral

700                                                  Agree

701                                                  Agree


    InstituteViews. Topic:11. If I had a workplace issue it was dealt w
ith quickly  \
0                                                    Agree

1                                                    Agree

2                                                    Agree
```

```
3                                                  Agree

4                                           Strongly Agree

..                                                    ...

697                                                Agree

698                                                  NaN

699                                                Agree

700                                       Strongly Disagree

701                                          Strongly Agree


    InstituteViews. Topic:12. If I had a workplace issue it was dealt w
ith efficiently  \
0                                                  Agree

1                                                  Agree

2                                                Neutral

3                                                  Agree

4                                                  Agree

..                                                    ...

697                                                Agree

698                                                  NaN

699                                                Agree

700                                       Strongly Disagree
```

```
701                                          Strongly Agree


    InstituteViews. Topic:13. If I had a workplace issue it was dealt w
ith discreetly  \
0                                                    Agree

1                                                 Disagree

2                                                  Neutral

3                                                    Agree

4                                           Strongly Agree

..                                                     ...

697                                                  Agree

698                                                    NaN

699                                                  Agree

700                                       Strongly Disagree

701                                          Strongly Agree


    WorkUnitViews. Topic:14. I was satisfied with the quality of the ma
nagement and supervision within my work unit  \
0                                                    Agree

1                                           Strongly Agree

2                                           Strongly Agree

3                                           Strongly Agree

4                                           Strongly Agree
```

```
..                                                  ...

697                                               Agree

698                                                 NaN

699                                               Agree

700                                    Strongly Disagree

701                                       Strongly Agree


    WorkUnitViews. Topic:15. I worked well with my colleagues  \
0                                                     Agree
1                                          Strongly Agree
2                                          Strongly Agree
3                                          Strongly Agree
4                                          Strongly Agree
..                                                    ...
697                                        Strongly Agree
698                                                 NaN
699                                               Agree
700                                               Agree
701                                        Strongly Agree

    WorkUnitViews. Topic:16. My job was challenging and interesting  \
0                                                     Agree
1                                          Strongly Agree
2                                                     Agree
3                                          Strongly Agree
4                                          Strongly Agree
..                                                    ...
697                                        Strongly Agree
698                                                 NaN
699                                        Strongly Agree
700                                        Strongly Agree
701                                               Neutral
```

```
    WorkUnitViews. Topic:17. I was encouraged to use my initiative in t
he course of my work  \
0                                          Strongly Agree

1                                                   Agree

2                                                   Agree

3                                          Strongly Agree

4                                          Strongly Agree

..                                                    ...

697                                        Strongly Agree

698                                                   NaN

699                                        Strongly Agree

700                                               Neutral

701                                        Strongly Agree


    WorkUnitViews. Topic:18. I had sufficient contact with other people
in my job  \
0                                                   Agree

1                                                   Agree

2                                                   Agree

3                                          Strongly Agree

4                                          Strongly Agree

..                                                    ...
```

```
697                                        Strongly Agree

698                                                  NaN

699                                                Agree

700                                                Agree

701                                        Strongly Agree


    WorkUnitViews. Topic:19. I was given adequate support and co-operat
ion by my peers to enable me to do my job  \
0                                                  Agree

1                                                  Agree

2                                                  Agree

3                                          Strongly Agree

4                                          Strongly Agree

..                                                   ...

697                                        Strongly Agree

698                                                  NaN

699                                                Agree

700                                             Disagree

701                                                Agree


    WorkUnitViews. Topic:20. I was able to use the full range of my ski
lls in my job  \
```

```
0                                                    Agree

1                                           Strongly Agree

2                                                    Agree

3                                           Strongly Agree

4                                           Strongly Agree

..                                                     ...

697                                                Neutral

698                                                    NaN

699                                                  Agree

700                                         Strongly Agree

701                                               Disagree


    WorkUnitViews. Topic:21. I was able to use the full range of my abi
lities in my job. ; Category:Level of Agreement; Question:YOUR VIEWS AB
OUT YOUR WORK UNIT]  \
0                                                    Agree

1                                                    Agree

2                                                    Agree

3                                           Strongly Agree

4                                           Strongly Agree
```

```
..                                                    ...

697                                                Neutral

698                                                    NaN

699                                                  Agree

700                                       Strongly Agree

701                                       Strongly Agree


    WorkUnitViews. Topic:22. I was able to use the full range of my kno
wledge in my job  \
0                                                    Agree

1                                                    Agree

2                                                    Agree

3                                         Strongly Agree

4                                         Strongly Agree

..                                                    ...

697                                                Neutral

698                                                    NaN
```

```
699                                                  Agree

700                                       Strongly Agree

701                                       Strongly Agree


    WorkUnitViews. Topic:23. My job provided sufficient variety  \
0                                                    Agree
1                                                    Agree
2                                                    Agree
3                                         Strongly Agree
4                                         Strongly Agree
..                                                     ...
697                                                  Agree
698                                                    NaN
699                                       Strongly Agree
700                                       Strongly Agree
701                                       Strongly Agree

    WorkUnitViews. Topic:24. I was able to cope with the level of stres
s and pressure in my job  \
0                                                    Agree

1                                         Strongly Agree

2                                                    Agree

3                                         Strongly Agree

4                                         Strongly Agree

..                                                     ...

697                                                  Agree

698                                                    NaN

699                                                  Agree
```

```
700                                  Strongly Disagree

701                                     Strongly Agree


    WorkUnitViews. Topic:25. My job allowed me to balance the demands o
f work and family to my satisfaction  \
0                                                 Agree

1                                                 Agree

2                                                 Agree

3                                     Strongly Agree

4                                     Strongly Agree

..                                                 ...

697                                          Neutral

698                                              NaN

699                                   Strongly Agree

700                                  Strongly Disagree

701                                     Strongly Agree


    WorkUnitViews. Topic:26. My supervisor gave me adequate personal re
cognition and feedback on my performance  \
0                                                 Agree

1                                     Strongly Agree

2                                                 Agree
```

```
3                                                Strongly Agree

4                                                Strongly Agree

..                                                          ...

697                                                     Neutral

698                                                         NaN

699                                              Strongly Agree

700                                                    Disagree

701                                              Strongly Agree


    WorkUnitViews. Topic:27. My working environment was satisfactory e.
g. sufficient space, good lighting, suitable seating and working area
 \
0                                                         Agree


1                                                Strongly Agree


2                                                         Agree


3                                                Strongly Agree


4                                                Strongly Agree


..                                                          ...


697                                                       Agree
```

```
698                                                  NaN

699                                       Strongly Agree

700                                                Agree

701                                       Strongly Agree


    WorkUnitViews. Topic:28. I was given the opportunity to mentor and
 coach others in order for me to pass on my skills and knowledge prior
 to my cessation date  \
0                                                Neutral

1                                                  Agree

2                                                  Agree

3                                         Strongly Agree

4                                         Strongly Agree

..                                                   ...

697                                              Neutral
```

```
698                                                  NaN

699                                                Agree

700                                       Strongly Agree

701                                              Neutral


    WorkUnitViews. Topic:29. There was adequate communication between s
taff in my unit  \
0                                                  Agree

1                                                  Agree

2                                                  Agree

3                                         Strongly Agree

4                                         Strongly Agree

..                                                   ...

697                                              Neutral

698                                                  NaN

699                                                Agree

700                                                Agree

701                                       Strongly Agree


    WorkUnitViews. Topic:30. Staff morale was positive within my work u
```

```
nit  \
0                                                  Agree

1                                         Strongly Agree

2                                                  Agree

3                                         Strongly Agree

4                                         Strongly Agree

..                                                   ...

697                                              Neutral

698                                                 NaN

699                                              Agree

700                                              Neutral

701                                       Strongly Agree


    Induction. Did you undertake Workplace Induction?  \
0                                                  Yes
1                                                   No
2                                                   No
3                                                  Yes
4                                                  Yes
..                                                 ...
697                                                 No
698                                                NaN
699                                                Yes
700                                                Yes
701                                                Yes

    InductionInfo. Topic:Did you undertake a Corporate Induction?  \
0                                                  Yes
```

```
1                                                    NaN
2                                                    NaN
3                                                     No
4                                                    Yes
..                                                   ...
697                                                  NaN
698                                                  NaN
699                                                   No
700                                                   No
701                                                  Yes

    InductionInfo. Topic:Did you undertake a Institute Induction?  \
0                                                    Yes
1                                                    NaN
2                                                    NaN
3                                                    Yes
4                                                    Yes
..                                                   ...
697                                                  NaN
698                                                  NaN
699                                                  Yes
700                                                  Yes
701                                                  Yes

    InductionInfo. Topic: Did you undertake Team Induction?  \
0                                                    Yes
1                                                    NaN
2                                                    NaN
3                                                    Yes
4                                                    Yes
..                                                   ...
697                                                  NaN
698                                                  NaN
699                                                  Yes
700                                                  Yes
701                                                  Yes

    InductionInfo. Face to Face Topic:Did you undertake a Corporate Ind
uction; Category:How it was conducted?  \
```

```
0                                                Face to Face

1                                                         NaN

2                                                         NaN

3                                                           -

4                                                           -

..                                                        ...

697                                                       NaN

698                                                       NaN

699                                                         -

700                                                         -

701                                                         -


    InductionInfo. On-line Topic:Did you undertake a Corporate Inductio
n; Category:How it was conducted?  \
0                                                           -

1                                                         NaN

2                                                         NaN

3                                                           -

4                                                           -

..                                                        ...

697                                                       NaN
```

```
698                                                   NaN

699                                                     -

700                                                On-line

701                                                     -


    InductionInfo. Induction Manual Topic:Did you undertake a Corporate
 Induction?  \
0                                                       -

1                                                     NaN

2                                                     NaN

3                                                       -

4                                          Induction Manual

..                                                    ...

697                                                   NaN

698                                                   NaN

699                                                     -

700                                                     -

701                                                     -


    InductionInfo. Face to Face Topic:Did you undertake a Institute Ind
uction?  \
0                                             Face to Face

1                                                     NaN
```

```
2                                                    NaN

3                                                    NaN

4                                           Face to Face

..                                                   ...

697                                                  NaN

698                                                  NaN

699                                                    -

700                                         Face to Face

701                                                    -


    InductionInfo. On-line Topic:Did you undertake a Institute Inductio
n?  \
0                                                      -

1                                                    NaN

2                                                    NaN

3                                                      -

4                                                      -

..                                                   ...

697                                                  NaN

698                                                  NaN

699                                                    -
```

```
700                                                    -

701                                                    -


    InductionInfo. Induction Manual Topic:Did you undertake a Institute
 Induction?  \
0                                                      -

1                                                    NaN

2                                                    NaN

3                                                      -

4                                                      -

..                                                   ...

697                                                  NaN

698                                                  NaN

699                                    Induction Manual

700                                                    -

701                                                    -


    InductionInfo. Face to Face Topic: Did you undertake Team Inductio
n; Category?  \
0                                          Face to Face

1                                                    NaN

2                                                    NaN
```

```
3                                                    -

4                                           Face to Face

..                                                 ...

697                                                NaN

698                                                NaN

699                                         Face to Face

700                                         Face to Face

701                                                  -


    InductionInfo. On-line Topic: Did you undertake Team Induction?proc
ess you undertook and how it was conducted.]  \
0                                                    -

1                                                  NaN

2                                                  NaN

3                                                    -

4                                                    -

..                                                 ...

697                                                NaN

698                                                NaN

699                                                  -

700                                                  -
```

```
701                                                   -


    InductionInfo. Induction Manual Topic: Did you undertake Team Induc
tion?  \
0                                                     -

1                                                   NaN

2                                                   NaN

3                                                     -

4                                                     -

..                                                  ...

697                                                 NaN

698                                                 NaN

699                                                   -

700                                                   -

701                                                   -


    Workplace. Topic:Did you and your Manager develop a Performance and
 Professional Development Plan (PPDP)?  \
0                                                   Yes

1                                                   Yes

2                                                   Yes

3                                                   Yes

4                                                   Yes
```

```
..                                                  ...
697                                                 Yes
698                                                 NaN
699                                                 Yes
700                                                  No
701                                                 Yes

    Workplace. Topic:Does your workplace promote a work culture free fr
om all forms of unlawful discrimination?  \
0                                                   Yes
1                                                   Yes
2                                                   Yes
3                                                   Yes
4                                                   Yes
..                                                  ...
697                                                 Yes
698                                                 NaN
699                                                 Yes
700                                                  No
701                                                 Yes
```

```
    Workplace. Topic:Does your workplace promote and practice the princ
iples of employment equity?  \
0                                                   Yes

1                                                   Yes

2                                                   Yes

3                                                   Yes

4                                                   Yes

..                                                  ...

697                                                 Yes

698                                                 NaN

699                                                 Yes

700                                                  No

701                                                 Yes


    Workplace. Topic:Does your workplace value the diversity of its emp
loyees?  \
0                                                   Yes

1                                                   Yes

2                                                   Yes

3                                                   Yes

4                                                   Yes

..                                                  ...
```

```
697                                                 Yes

698                                                 NaN

699                                                 Yes

700                                                 Yes

701                                                 Yes


    Workplace. Topic:Would you recommend the Institute as an employer t
o others?  \
0                                                   Yes

1                                                   Yes

2                                                   Yes

3                                                   Yes

4                                                   Yes

..                                                  ...

697                                                 Yes

698                                                 NaN

699                                                 Yes

700                                                  No

701                                                 Yes


    Gender.     What is your Gender? CurrentAge.     Current Age  \
0                                Female                    26 – 30
1                                   NaN                        NaN
```

```
2                                  NaN                          NaN
3                                  NaN                          NaN
4                                 Male                      41 – 45
..                                 ...                          ...
697                               Male                        51-55
698                                NaN                          NaN
699                             Female                        51-55
700                             Female                      41 – 45
701                             Female                      26 – 30

    Employment Type.     Employment Type Classification.     Classifica
tion  \
0                    Temporary Full-time                Administration
 (AO)
1                                    NaN
 NaN
2                                    NaN
 NaN
3                                    NaN
 NaN
4                    Permanent Full-time            Teacher (including
 LVT)
..                                   ...
 ...
697                  Temporary Full-time            Teacher (including
 LVT)
698                                  NaN
 NaN
699                  Permanent Full-time            Teacher (including
 LVT)
700                  Temporary Full-time          Professional Officer
 (PO)
701                      Contract/casual                Administration
 (AO)

    LengthofServiceOverall. Overall Length of Service at Institute (in
 years)  \
0                                                  1-2
```

```
1                                                    NaN

2                                                    NaN

3                                                    NaN

4                                                    3-4

..                                                   ...

697                                                  1-2

698                                                  NaN

699                                                  5-6

700                                                  1-2

701                                                  3-4


    LengthofServiceCurrent. Length of Service at current workplace (in
 years)
0                                                    1-2

1                                                    NaN

2                                                    NaN

3                                                    NaN

4                                                    3-4

..                                                   ...

697                                                  1-2

698                                                  NaN
```

```
699                                                  1-2

700                                                  1-2

701                                                  1-2


[702 rows x 72 columns]>
```

In [3]:
```
# Read the dete_survey.csv CSV file into pandas, and assign it to the v
ariable name dete_survey.
dete_survey = pd.read_csv('dete-exit-survey-january-2014.csv', encoding
= "cp1252")
#Quick exploration of the data.
dete_survey.head()
```

Out[3]:

	ID	SeparationType	Cease Date	DETE Start Date	Role Start Date	Position	Classification	Region	Busi
0	1	Ill Health Retirement	08/2012	1984	2004	Public Servant	A01-A04	Central Office	Corp Stra Peform
1	2	Voluntary Early Retirement (VER)	08/2012	Not Stated	Not Stated	Public Servant	AO5-AO7	Central Office	Corp Stra Peform
2	3	Voluntary Early Retirement (VER)	05/2012	2011	2011	Schools Officer	NaN	Central Office	Educ Queen
3	4	Resignation-Other reasons	05/2012	2005	2006	Teacher	Primary	Central Queensland	
4	5	Age Retirement	05/2012	1970	1989	Head of Curriculum/Head of Special Education	NaN	South East	

In [4]:
```
dete_survey.info
```

Out[4]:
```
<bound method DataFrame.info of       ID                         Separat
ionType Cease Date DETE Start Date  \
0      1                 Ill Health Retirement    08/2012            19
84
1      2      Voluntary Early Retirement (VER)    08/2012      Not Stat
ed
2      3      Voluntary Early Retirement (VER)    05/2012            20
11
3      4              Resignation-Other reasons    05/2012            20
05
4      5                         Age Retirement    05/2012            19
70
..   ...                                   ...        ...
...
817  819                         Age Retirement    02/2014            19
77
818  820                         Age Retirement    01/2014            19
80
819  821  Resignation-Move overseas/interstate    01/2014            20
09
820  822                  Ill Health Retirement    12/2013            20
01
821  823  Resignation-Move overseas/interstate    12/2013      Not Stat
ed

    Role Start Date                                    Position  \
0              2004                                Public Servant
1        Not Stated                                Public Servant
2              2011                               Schools Officer
3              2006                                       Teacher
4              1989  Head of Curriculum/Head of Special Education
..              ...                                           ...
817            1999                                       Teacher
818            1980                                       Teacher
819            2009                                Public Servant
820            2009                                       Teacher
```

```
821      Not Stated                                  Teacher Aide

    Classification                    Region  \
0          A01-A04            Central Office
1          AO5-AO7            Central Office
2              NaN            Central Office
3          Primary         Central Queensland
4              NaN                 South East
..             ...                        ...
817        Primary         Central Queensland
818      Secondary                North Coast
819        A01-A04            Central Office
820      Secondary  Darling Downs South West
821            NaN              Metropolitan

                         Business Unit    Employment Status  \
0    Corporate Strategy and Peformance  Permanent Full-time
1    Corporate Strategy and Peformance  Permanent Full-time
2                Education Queensland  Permanent Full-time
3                                 NaN  Permanent Full-time
4                                 NaN  Permanent Full-time
..                                ...                  ...
817                               NaN  Permanent Part-time
818                               NaN  Permanent Full-time
819              Education Queensland  Permanent Full-time
820                               NaN  Permanent Full-time
821                               NaN                  NaN

     Career move to public sector  Career move to private sector  \
0                            True                          False
1                           False                          False
2                           False                          False
3                           False                           True
4                           False                          False
..                            ...                            ...
817                         False                          False
818                         False                          False
819                         False                          False
820                         False                          False
```

```
821                          False                          False

     Interpersonal conflicts  Job dissatisfaction  \
0                      False                 True
1                      False                False
2                      False                False
3                      False                False
4                      False                False
..                       ...                  ...
817                    False                False
818                    False                False
819                    False                False
820                    False                False
821                    False                False

     Dissatisfaction with the department  Physical work environment  \
0                                  False                       False
1                                  False                       False
2                                  False                       False
3                                  False                       False
4                                  False                       False
..                                   ...                         ...
817                                False                       False
818                                False                       False
819                                False                       False
820                                False                       False
821                                False                       False

     Lack of recognition  Lack of job security  Work location  \
0                   True                 False          False
1                  False                 False          False
2                  False                 False          False
3                  False                 False          False
4                  False                 False          False
..                   ...                   ...            ...
817                False                 False          False
818                False                 False          False
819                False                 False          False
820                False                 False          False
```

```
821                False                False         False

     Employment conditions  Maternity/family  Relocation  Study/Travel
\
0                    False             False       False         False

1                    False             False       False         False

2                    False             False       False         False

3                    False             False       False         False

4                    False             False       False         False

..                     ...               ...         ...           ...

817                  False              True       False         False

818                  False             False       False         False

819                  False              True        True         False

820                  False             False       False         False

821                  False             False       False         False


     Ill Health  Traumatic incident  Work life balance  Workload  \
0         False               False              False     False
1         False               False              False     False
2         False               False              False     False
3         False               False              False     False
4         False               False               True     False
..          ...                 ...                ...       ...
817        True               False              False      True
818       False               False              False     False
819       False               False               True     False
820        True               False              False     False
821       False               False              False     False
```

```
     None of the above Professional Development Opportunities for promo
tion  \
0                  True                          A
   A
1                 False                          A
   A
2                  True                          N
   N
3                 False                          A
   N
4                 False                          A
   A
..                  ...                        ...
...
817               False                         SA
   N
818                True                          A
  SA
819               False                          A
   A
820               False                          A
   D
821               False                        NaN
NaN

    Staff morale Workplace issue Physical environment Worklife balance
\
0              N               N                    N                A

1              N               N                    N                N

2              N               N                    N                N

3              N               N                    A                A

4              N               N                    D                D

..           ...             ...                  ...              ...
```

```
817            D              D                   A              N
818            D              D                   D              A
819            A              A                   A              D
820            D              A                  SD             SD
821          NaN            NaN                 NaN            NaN

    Stress and pressure support Performance of supervisor Peer support
\
0                            A                      A           A
1                            A                      A           A
2                            N                      N           N
3                            N                      N           A
4                            N                      A           A
..                         ...                    ...         ...
817                          N                      D           A
818                          A                      N           A
819                          N                      A           A
820                         SD                      A           D
821                        NaN                    NaN         NaN

    Initiative Skills Coach Career Aspirations Feedback Further PD  \
0            N      N     N                   A        A          A
```

```
1           N      N     N                A        A          A
2           N      N     N                N        N          N
3           A      A     A                A        A          A
4           A      A     A                A       SA         SA
..        ...    ...   ...              ...      ...        ...
817         N      A     A                N       SA         SA
818         N      A     A                N        A          N
819         A      A     A                A        A          A
820        SD     SD     D                A        A          N
821       NaN    NaN   NaN              NaN      NaN        NaN

    Communication My say Information Kept informed Wellness programs  \
0               N      A           A              N                  N
1               N      A           A              N                  N
2               A      A           N              N                  N
3               A      A           A              A                  N
4               D      D           A              N                  A
..            ...    ...         ...            ...                ...
817             N      D           A              A                  A
818             N      A           A              N                  N
819             A      A           A              A                  N
820             N      N          SD              A                  N
821           NaN    NaN         NaN            NaN                NaN

    Health & Safety  Gender          Age Aboriginal Torres Strait South
Sea  \
0                 N    Male        56-60        NaN           NaN
NaN
1                 N    Male        56-60        NaN           NaN
NaN
2                 N    Male  61 or older        NaN           NaN
NaN
3                 A  Female        36-40        NaN           NaN
NaN
4                 M  Female  61 or older        NaN           NaN
NaN
..              ...     ...          ...        ...           ...
...
817              SA  Female        56-60        NaN           NaN
```

```
NaN
818               N    Male        51-55        NaN           NaN
NaN
819               A  Female        31-35        NaN           NaN
NaN
820               A  Female        41-45        NaN           NaN
NaN
821             NaN     NaN          NaN        NaN           NaN
NaN

    Disability NESB
0          NaN  Yes
1          NaN  NaN
2          NaN  NaN
3          NaN  NaN
4          NaN  NaN
..         ...  ...
817        NaN  NaN
818        NaN  NaN
819        NaN  NaN
820        NaN  NaN
821        NaN  NaN

[822 rows x 56 columns]>
```

In [5]:
```
# Check out missing values.
dete_missing = dete_survey.isnull().sum()
dete_missing
```

Out[5]:
```
ID                                     0
SeparationType                         0
Cease Date                             0
DETE Start Date                        0
Role Start Date                        0
Position                               5
Classification                       367
Region                                 0
Business Unit                        696
Employment Status                      5
```

Career move to public sector	0
Career move to private sector	0
Interpersonal conflicts	0
Job dissatisfaction	0
Dissatisfaction with the department	0
Physical work environment	0
Lack of recognition	0
Lack of job security	0
Work location	0
Employment conditions	0
Maternity/family	0
Relocation	0
Study/Travel	0
Ill Health	0
Traumatic incident	0
Work life balance	0
Workload	0
None of the above	0
Professional Development	14
Opportunities for promotion	87
Staff morale	6
Workplace issue	34
Physical environment	5
Worklife balance	7
Stress and pressure support	12
Performance of supervisor	9
Peer support	10
Initiative	9
Skills	11
Coach	55
Career Aspirations	76
Feedback	30
Further PD	54
Communication	8
My say	10
Information	6
Kept informed	9
Wellness programs	56
Health & Safety	29

```
Gender                                 24
Age                                    11
Aboriginal                            806
Torres Strait                         819
South Sea                             815
Disability                            799
NESB                                  790
dtype: int64
```

In [6]:
```
#  Check out missing value.
tafe_missing = tafe_survey.isnull().sum()
tafe_missing
```

Out[6]:
```
Record ID
        0
Institute
        0
WorkArea
        0
CESSATION YEAR
        7
Reason for ceasing employment
        1

     ...
CurrentAge.     Current Age
      106
Employment Type.     Employment Type
      106
Classification.     Classification
      106
LengthofServiceOverall. Overall Length of Service at Institute (in year
s)    106
LengthofServiceCurrent. Length of Service at current workplace (in year
s)    106
Length: 72, dtype: int64
```

In [7]:
```
tafe_survey['Reason for ceasing employment'].value_counts()
```

```
Out[7]: Resignation                 340
        Contract Expired            127
        Retrenchment/ Redundancy    104
        Retirement                   82
        Transfer                     25
        Termination                  23
        Name: Reason for ceasing employment, dtype: int64
```

We can make the following observations based on the work above: The dete_survey dataframe contains 'Not Stated' values that indicate values are missing, but they aren't represented as NaN. Both the dete_survey and tafe_survey contain many columns that we don't need to complete our analysis. Each dataframe contains many of the same columns, but the column names are different. There are multiple columns/answers that indicate an employee resigned because they were dissatisfied.

Exploring the missinig values and droping unimportant columns.

First, we'll correct the Not Stated values and drop some of the columns we don't need for our analysis.

```
In [8]: dete_survey = pd.read_csv('dete-exit-survey-january-2014.csv', encoding
        = "cp1252",na_values = 'Not Stated')
        tafe_survey = pd.read_csv('tafe-employee-exit-survey-access-database-de
        cember-2013.csv',encoding='cp1252', na_values='Not Stated')
```

```
In [9]: tafe_survey.columns[17:66]
```

```
Out[9]: Index(['Main Factor.     Which of these was the main factor for leavin
        g?',
               'InstituteViews. Topic:1. I feel the senior leadership had a cle
        ar vision and direction',
               'InstituteViews. Topic:2. I was given access to skills training
        to help me do my job better',
               'InstituteViews. Topic:3. I was given adequate opportunities for
        personal development',
               'InstituteViews. Topic:4. I was given adequate opportunities for
        promotion within %Institute]Q25LBL%',
               'InstituteViews. Topic:5. I felt the salary for the job was righ
```

```
t for the responsibilities I had',
       'InstituteViews. Topic:6. The organisation recognised when staff
did good work',
       'InstituteViews. Topic:7. Management was generally supportive of
me',
       'InstituteViews. Topic:8. Management was generally supportive of
my team',
       'InstituteViews. Topic:9. I was kept informed of the changes in
the organisation which would affect me',
       'InstituteViews. Topic:10. Staff morale was positive within the
Institute',
       'InstituteViews. Topic:11. If I had a workplace issue it was dea
lt with quickly',
       'InstituteViews. Topic:12. If I had a workplace issue it was dea
lt with efficiently',
       'InstituteViews. Topic:13. If I had a workplace issue it was dea
lt with discreetly',
       'WorkUnitViews. Topic:14. I was satisfied with the quality of th
e management and supervision within my work unit',
       'WorkUnitViews. Topic:15. I worked well with my colleagues',
       'WorkUnitViews. Topic:16. My job was challenging and interestin
g',
       'WorkUnitViews. Topic:17. I was encouraged to use my initiative
in the course of my work',
       'WorkUnitViews. Topic:18. I had sufficient contact with other pe
ople in my job',
       'WorkUnitViews. Topic:19. I was given adequate support and co-op
eration by my peers to enable me to do my job',
       'WorkUnitViews. Topic:20. I was able to use the full range of my
skills in my job',
       'WorkUnitViews. Topic:21. I was able to use the full range of my
abilities in my job. ; Category:Level of Agreement; Question:YOUR VIEWS
ABOUT YOUR WORK UNIT]',
       'WorkUnitViews. Topic:22. I was able to use the full range of my
knowledge in my job',
       'WorkUnitViews. Topic:23. My job provided sufficient variety',
       'WorkUnitViews. Topic:24. I was able to cope with the level of s
tress and pressure in my job',
       'WorkUnitViews. Topic:25. My job allowed me to balance the deman
```

```
ds of work and family to my satisfaction',
       'WorkUnitViews. Topic:26. My supervisor gave me adequate persona
l recognition and feedback on my performance',
       'WorkUnitViews. Topic:27. My working environment was satisfactor
y e.g. sufficient space, good lighting, suitable seating and working ar
ea',
       'WorkUnitViews. Topic:28. I was given the opportunity to mentor
and coach others in order for me to pass on my skills and knowledge pri
or to my cessation date',
       'WorkUnitViews. Topic:29. There was adequate communication betwe
en staff in my unit',
       'WorkUnitViews. Topic:30. Staff morale was positive within my wo
rk unit',
       'Induction. Did you undertake Workplace Induction?',
       'InductionInfo. Topic:Did you undertake a Corporate Induction?',
       'InductionInfo. Topic:Did you undertake a Institute Induction?',
       'InductionInfo. Topic: Did you undertake Team Induction?',
       'InductionInfo. Face to Face Topic:Did you undertake a Corporate
Induction; Category:How it was conducted?',
       'InductionInfo. On-line Topic:Did you undertake a Corporate Indu
ction; Category:How it was conducted?',
       'InductionInfo. Induction Manual Topic:Did you undertake a Corpo
rate Induction?',
       'InductionInfo. Face to Face Topic:Did you undertake a Institute
Induction?',
       'InductionInfo. On-line Topic:Did you undertake a Institute Indu
ction?',
       'InductionInfo. Induction Manual Topic:Did you undertake a Insti
tute Induction?',
       'InductionInfo. Face to Face Topic: Did you undertake Team Induc
tion; Category?',
       'InductionInfo. On-line Topic: Did you undertake Team Induction?
process you undertook and how it was conducted.]',
       'InductionInfo. Induction Manual Topic: Did you undertake Team I
nduction?',
       'Workplace. Topic:Did you and your Manager develop a Performance
and Professional Development Plan (PPDP)?',
       'Workplace. Topic:Does your workplace promote a work culture fre
e from all forms of unlawful discrimination?',
```

```
       'Workplace. Topic:Does your workplace promote and practice the p
rinciples of employment equity?',
       'Workplace. Topic:Does your workplace value the diversity of its
employees?',
       'Workplace. Topic:Would you recommend the Institute as an employ
er to others?'],
      dtype='object')
```

In [10]:
```
# droping these above columns.

tafe_survey_updated = tafe_survey.drop(tafe_survey.columns[17:66], axis
=1)
```

In [11]:
```
dete_survey.columns[28:49]
```

Out[11]:
```
Index(['Professional Development', 'Opportunities for promotion',
       'Staff morale', 'Workplace issue', 'Physical environment',
       'Worklife balance', 'Stress and pressure support',
       'Performance of supervisor', 'Peer support', 'Initiative', 'Skil
ls',
       'Coach', 'Career Aspirations', 'Feedback', 'Further PD',
       'Communication', 'My say', 'Information', 'Kept informed',
       'Wellness programs', 'Health & Safety'],
      dtype='object')
```

In [12]:
```
# Droping theses columns.
dete_survey_updated = dete_survey.drop(dete_survey.columns[28:49],axis=
1)
```

In [13]:
```
print(dete_survey_updated)
print('\n')
print(tafe_survey_updated)
```

```
      ID                        SeparationType Cease Date  DETE Start D
ate  \
0      1                 Ill Health Retirement    08/2012           198
4.0
1      2      Voluntary Early Retirement (VER)    08/2012
```

```
NaN
2      3      Voluntary Early Retirement (VER)    05/2012           201
1.0
3      4             Resignation-Other reasons    05/2012           200
5.0
4      5                       Age Retirement    05/2012           197
0.0
..   ...                                 ...        ...
...
817  819                       Age Retirement    02/2014           197
7.0
818  820                       Age Retirement    01/2014           198
0.0
819  821  Resignation-Move overseas/interstate    01/2014           200
9.0
820  822                 Ill Health Retirement    12/2013           200
1.0
821  823  Resignation-Move overseas/interstate    12/2013
NaN

     Role Start Date                                    Position  \
0             2004.0                              Public Servant
1                NaN                              Public Servant
2             2011.0                             Schools Officer
3             2006.0                                     Teacher
4             1989.0  Head of Curriculum/Head of Special Education
..               ...                                         ...
817           1999.0                                     Teacher
818           1980.0                                     Teacher
819           2009.0                              Public Servant
820           2009.0                                     Teacher
821              NaN                                Teacher Aide

    Classification                    Region  \
0          A01-A04           Central Office
1          AO5-AO7           Central Office
2              NaN           Central Office
3          Primary       Central Queensland
4              NaN               South East
```

```
..            ...                        ...
817       Primary       Central Queensland
818     Secondary              North Coast
819       A01-A04           Central Office
820     Secondary  Darling Downs South West
821           NaN             Metropolitan

                         Business Unit    Employment Status  \
0    Corporate Strategy and Peformance  Permanent Full-time
1    Corporate Strategy and Peformance  Permanent Full-time
2                Education Queensland  Permanent Full-time
3                                 NaN  Permanent Full-time
4                                 NaN  Permanent Full-time
..                                ...                  ...
817                               NaN  Permanent Part-time
818                               NaN  Permanent Full-time
819              Education Queensland  Permanent Full-time
820                               NaN  Permanent Full-time
821                               NaN                  NaN

     Career move to public sector  Career move to private sector  \
0                            True                          False
1                           False                          False
2                           False                          False
3                           False                           True
4                           False                          False
..                            ...                            ...
817                         False                          False
818                         False                          False
819                         False                          False
820                         False                          False
821                         False                          False

     Interpersonal conflicts  Job dissatisfaction  \
0                      False                 True
1                      False                False
2                      False                False
3                      False                False
4                      False                False
```

```
..                      ...                 ...
817                   False               False
818                   False               False
819                   False               False
820                   False               False
821                   False               False

     Dissatisfaction with the department  Physical work environment  \
0                                  False                       False
1                                  False                       False
2                                  False                       False
3                                  False                       False
4                                  False                       False
..                                   ...                         ...
817                                False                       False
818                                False                       False
819                                False                       False
820                                False                       False
821                                False                       False

     Lack of recognition  Lack of job security  Work location  \
0                   True                 False          False
1                  False                 False          False
2                  False                 False          False
3                  False                 False          False
4                  False                 False          False
..                   ...                   ...            ...
817                False                 False          False
818                False                 False          False
819                False                 False          False
820                False                 False          False
821                False                 False          False

     Employment conditions  Maternity/family  Relocation  Study/Travel
\
0                    False             False       False         False

1                    False             False       False         False
```

2	False	False	False	False
3	False	False	False	False
4	False	False	False	False
..	...	...	...	...
817	False	True	False	False
818	False	False	False	False
819	False	True	True	False
820	False	False	False	False
821	False	False	False	False

	Ill Health	Traumatic incident	Work life balance	Workload
0	False	False	False	False
1	False	False	False	False
2	False	False	False	False
3	False	False	False	False
4	False	False	True	False
..	...	...	...	...
817	True	False	False	True
818	False	False	False	False
819	False	False	True	False
820	True	False	False	False
821	False	False	False	False

	None of the above	Gender	Age	Aboriginal	Torres Strait
0	True	Male	56-60	NaN	NaN
1	False	Male	56-60	NaN	NaN
2	True	Male	61 or older	NaN	NaN
3	False	Female	36-40	NaN	NaN
4	False	Female	61 or older	NaN	NaN
..	...	...	...	...	...

```
817              False  Female        56-60        NaN           NaN
818               True    Male        51-55        NaN           NaN
819              False  Female        31-35        NaN           NaN
820              False  Female        41-45        NaN           NaN
821              False     NaN          NaN        NaN           NaN

    South Sea Disability NESB
0         NaN        NaN  Yes
1         NaN        NaN  NaN
2         NaN        NaN  NaN
3         NaN        NaN  NaN
4         NaN        NaN  NaN
..        ...        ...  ...
817       NaN        NaN  NaN
818       NaN        NaN  NaN
819       NaN        NaN  NaN
820       NaN        NaN  NaN
821       NaN        NaN  NaN

[822 rows x 35 columns]


              Record ID                             Institute  \
0    634133009996094000  Southern Queensland Institute of TAFE
1    634133654064531000            Mount Isa Institute of TAFE
2    634138845606563000            Mount Isa Institute of TAFE
3    634139903350000000            Mount Isa Institute of TAFE
4    634146578511788000  Southern Queensland Institute of TAFE
..                  ...                                    ...
697  635066785175197000         Barrier Reef Institute of TAFE
698  635067716405666000  Southern Queensland Institute of TAFE
699  635070442972541000       Tropical North Institute of TAFE
700  635071153170979000      Southbank Institute of Technology
701  635073030973791000       Tropical North Institute of TAFE

                    WorkArea  CESSATION YEAR Reason for ceasing employ
ment  \
0    Non-Delivery (corporate)          2010.0              Contract Exp
ired
```

```
1    Non-Delivery (corporate)          2010.0                    Retire
ment
2         Delivery (teaching)          2010.0                    Retire
ment
3    Non-Delivery (corporate)          2010.0                   Resigna
tion
4         Delivery (teaching)          2010.0                   Resigna
tion
..                        ...             ...
...
697       Delivery (teaching)          2013.0                   Resigna
tion
698  Non-Delivery (corporate)          2013.0                   Resigna
tion
699       Delivery (teaching)          2013.0                   Resigna
tion
700  Non-Delivery (corporate)          2013.0              Contract Exp
ired
701  Non-Delivery (corporate)          2013.0                   Resigna
tion

    Contributing Factors. Career Move - Public Sector   \
0                                                  NaN
1                                                    -
2                                                    -
3                                                    -
4                                                    -
..                                                 ...
697                       Career Move - Public Sector
698                       Career Move - Public Sector
699                                                  -
700                                                NaN
701                                                  -

    Contributing Factors. Career Move - Private Sector   \
0                                                  NaN
1                                                    -
2                                                    -
3                                                    -
```

```
4                          Career Move - Private Sector
..                                                  ...
697                                                   -
698                                                   -
699                                                   -
700                                                 NaN
701                                                   -

    Contributing Factors. Career Move - Self-employment  \
0                                                   NaN
1                                                     -
2                                                     -
3                                                     -
4                                                     -
..                                                  ...
697                                                   -
698                                                   -
699                                                   -
700                                                 NaN
701                       Career Move - Self-employment

    Contributing Factors. Ill Health Contributing Factors. Maternity/Fa
mily  \
0                                 NaN
NaN
1                                   -
   -
2                                   -
   -
3                                   -
   -
4                                   -
   -
..                                ...
...
697                                 -
   -
698                                 -
   -
```

```
699                                    -
   -
700                                  NaN
NaN
701                                    -
   -

    Contributing Factors. Dissatisfaction  \
0                                      NaN
1                                        -
2                                        -
3                                        -
4                                        -
..                                     ...
697                                      -
698                                      -
699                                      -
700                                    NaN
701                                      -

    Contributing Factors. Job Dissatisfaction  \
0                                        NaN
1                                          -
2                                          -
3                                          -
4                                          -
..                                       ...
697                                        -
698                                        -
699                                        -
700                                      NaN
701                                        -

    Contributing Factors. Interpersonal Conflict Contributing Factors.
Study  \
0                                           NaN
  NaN
1                                             -
    -

2
```

```
2                                                    -
    -
3                                                    -
    -
4                                                    -
    -
..                                                 ...
  ...
697                                                  -
    -
698                                                  -
    -
699                                                  -
    -
700                                                NaN
  NaN
701                                                  -
    -

    Contributing Factors. Travel Contributing Factors. Other  \
0                            NaN                           NaN
1                         Travel                             -
2                              -                             -
3                         Travel                             -
4                              -                             -
..                           ...                           ...
697                            -                             -
698                            -                             -
699                            -                         Other
700                          NaN                           NaN
701                       Travel                             -

    Contributing Factors. NONE Gender.     What is your Gender?  \
0                             NaN                          Female
1                               -                             NaN
2                            NONE                             NaN
3                               -                             NaN
4                               -                            Male
..                            ...                             ...

                                                                 M
```

```
697                          -                          Male
698                          -                           NaN
699                          -                        Female
700                        NaN                        Female
701                          -                        Female

    CurrentAge.     Current Age Employment Type.     Employment Type  \
0                       26 – 30                  Temporary Full-time
1                           NaN                                  NaN
2                           NaN                                  NaN
3                           NaN                                  NaN
4                       41 – 45                  Permanent Full-time
..                          ...                                  ...
697                       51-55                  Temporary Full-time
698                         NaN                                  NaN
699                       51-55                  Permanent Full-time
700                     41 – 45                  Temporary Full-time
701                     26 – 30                      Contract/casual

    Classification.     Classification  \
0                  Administration (AO)
1                                  NaN
2                                  NaN
3                                  NaN
4              Teacher (including LVT)
..                                 ...
697            Teacher (including LVT)
698                                NaN
699            Teacher (including LVT)
700          Professional Officer (PO)
701                Administration (AO)

    LengthofServiceOverall. Overall Length of Service at Institute (in
years)  \
0                                                  1-2

1                                                  NaN

2                                                  NaN
```

```
3                                                    NaN

4                                                    3-4

..                                                   ...

697                                                  1-2

698                                                  NaN

699                                                  5-6

700                                                  1-2

701                                                  3-4


    LengthofServiceCurrent. Length of Service at current workplace (in
years)
0                                                    1-2

1                                                    NaN

2                                                    NaN

3                                                    NaN

4                                                    3-4

..                                                   ...

697                                                  1-2

698                                                  NaN

699                                                  1-2

700                                                  1-2
```

```
701                                                 1-2


[702 rows x 23 columns]
```

Rename Columns

Rename the same columes value of two dataframes which are: ID and Record ID , SeparationType Reason for ceasing employment, Cease Date CESSATION YEAR, DETE Start Date Empty, Empty LengthofServiceOverall.Overall Length of Service at Institute (in years) , Age CurrentAge.Current Age , Gender Gender.What is your Gender?

Next, we'll standardize the names of the columns we want to work with, because we eventually want to combine the dataframes.

In [14]:
```
# Clean the column names
dete_survey_updated.columns = dete_survey_updated.columns.str.lower().s
tr.strip().str.replace(' ', '_')

# Check that the column names were updated correctly
dete_survey_updated.columns
```

Out[14]:
```
Index(['id', 'separationtype', 'cease_date', 'dete_start_date',
       'role_start_date', 'position', 'classification', 'region',
       'business_unit', 'employment_status', 'career_move_to_public_sec
tor',
       'career_move_to_private_sector', 'interpersonal_conflicts',
       'job_dissatisfaction', 'dissatisfaction_with_the_department',
       'physical_work_environment', 'lack_of_recognition',
       'lack_of_job_security', 'work_location', 'employment_condition
s',
       'maternity/family', 'relocation', 'study/travel', 'ill_health',
       'traumatic_incident', 'work_life_balance', 'workload',
       'none_of_the_above', 'gender', 'age', 'aboriginal', 'torres_stra
it',
       'south_sea', 'disability', 'nesb'],
      dtype='object')
```

In [15]:
```
# Update column names to match the names in dete_survey_updated.
mapping={
    'Record ID': 'id',
    'CESSATION YEAR': 'cease_date',
    'Reason for ceasing employment': 'separationtype',
    'Gender. What is your Gender?': 'gender',
    'CurrentAge. Current Age': 'age',
    'Employment Type. Employment Type': 'employment_status',
    'Classification. Classification': 'position',
    'LengthofServiceOverall. Overall Length of Service at Institute (in
 years)': 'institute_service',
    'LengthofServiceCurrent. Length of Service at current workplace (in
 years)': 'role_service'}

tafe_survey_updated = tafe_survey_updated.rename(mapping, axis=1)
# Check that the specified column names were updated correctly.
tafe_survey_updated.columns
```

Out[15]:
```
Index(['id', 'Institute', 'WorkArea', 'cease_date', 'separationtype',
       'Contributing Factors. Career Move - Public Sector ',
       'Contributing Factors. Career Move - Private Sector ',
       'Contributing Factors. Career Move - Self-employment',
       'Contributing Factors. Ill Health',
       'Contributing Factors. Maternity/Family',
       'Contributing Factors. Dissatisfaction',
       'Contributing Factors. Job Dissatisfaction',
       'Contributing Factors. Interpersonal Conflict',
       'Contributing Factors. Study', 'Contributing Factors. Travel',
       'Contributing Factors. Other', 'Contributing Factors. NONE',
       'Gender.     What is your Gender?', 'CurrentAge.     Current Ag
e',
       'Employment Type.     Employment Type',
       'Classification.     Classification', 'institute_service',
       'role_service'],
      dtype='object')
```

Filtering of the data.

To answer for first question in this project, we'll only analyze survey respondents who resigned, so we'll only select separation types containing the string 'Resignation'.

In [16]:
```
# Check the unique values for the separationtype column.

tafe_survey_updated['separationtype'].value_counts()
```

Out[16]:
```
Resignation                 340
Contract Expired            127
Retrenchment/ Redundancy    104
Retirement                   82
Transfer                     25
Termination                  23
Name: separationtype, dtype: int64
```

In [17]:
```
dete_survey_updated['separationtype'].value_counts()
```

Out[17]:
```
Age Retirement                          285
Resignation-Other reasons               150
Resignation-Other employer               91
Resignation-Move overseas/interstate     70
Voluntary Early Retirement (VER)         67
Ill Health Retirement                    61
Other                                    49
Contract Expired                         34
Termination                              15
Name: separationtype, dtype: int64
```

In [18]:
```
# Select all Resignation  separation types of dataframe dete_survey_upd
ated.
dete_survey_updated['separationtype'] = dete_survey_updated['separation
type'].str.split('-').str[0]

# Check the values in the separationtype column were updated correctly
dete_survey_updated['separationtype'].value_counts()
```

Out[18]:
```
Resignation                 311
Age Retirement              285
Voluntary Early Retirement (VER)     67
```

```
Ill Health Retirement                 61
Other                                 49
Contract Expired                      34
Termination                           15
Name: separationtype, dtype: int64
```

In [19]:
```
# Select only the resignation separation types from each dataframe.

dete_resignations = dete_survey_updated[dete_survey_updated['separation
type'] == 'Resignation'].copy()
tafe_resignations = tafe_survey_updated[tafe_survey_updated['separation
type'] == 'Resignation'].copy()
```

Verifying Data.

In this step, we'll focus on verifying that the years in the cease_date and dete_start_date columns make sense. . Since the cease_date is the last year of the person's employment and the dete_start_date is the person's first year of employment, it wouldn't make sense to have years after the current date. . Given that most people in this field start working in their 20s, it's also unlikely that the dete_start_date was before the year 1940.

In [20]:
```
# Check the unique values.
dete_resignations['cease_date'].value_counts()
```

Out[20]:
```
2012       126
2013        74
01/2014     22
12/2013     17
06/2013     14
09/2013     11
11/2013      9
07/2013      9
10/2013      6
08/2013      4
05/2012      2
05/2013      2
07/2012      1
09/2010      1
```

```
2010         1
07/2006      1
Name: cease_date, dtype: int64
```

In [21]:
```
# Extract the years and convert them to a float type

dete_resignations['cease_date'] = dete_resignations['cease_date'].str.s
plit('/').str[-1]
dete_resignations['cease_date'] = dete_resignations['cease_date'].astyp
e(float)

# Check the values again and look for outliers.
dete_resignations['cease_date'].value_counts()
```

Out[21]:
```
2013.0    146
2012.0    129
2014.0     22
2010.0      2
2006.0      1
Name: cease_date, dtype: int64
```

In [22]:
```
# Check the unique values.
dete_resignations['dete_start_date'].value_counts().sort_values()
```

Out[22]:
```
1963.0     1
1971.0     1
1972.0     1
1984.0     1
1977.0     1
1987.0     1
1975.0     1
1973.0     1
1982.0     1
1974.0     2
1983.0     2
1976.0     2
1986.0     3
1985.0     3
2001.0     3
```

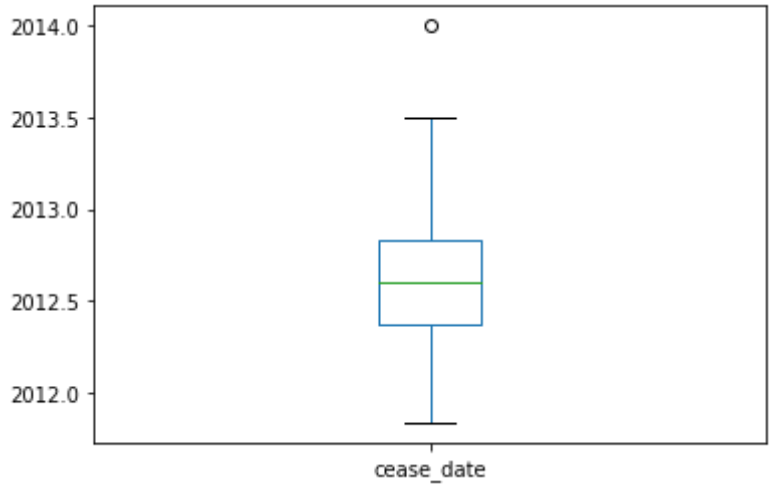
```
1995.0     4
1988.0     4
1989.0     4
1991.0     4
1997.0     5
1980.0     5
1993.0     5
1990.0     5
1994.0     6
2003.0     6
1998.0     6
1992.0     6
2002.0     6
1996.0     6
1999.0     8
2000.0     9
2013.0    10
2009.0    13
2006.0    13
2004.0    14
2005.0    15
2010.0    17
2012.0    21
2007.0    21
2008.0    22
2011.0    24
Name: dete_start_date, dtype: int64
```

In [23]:
```
# Check the unique values

tafe_resignations['cease_date'].value_counts().sort_values()
```

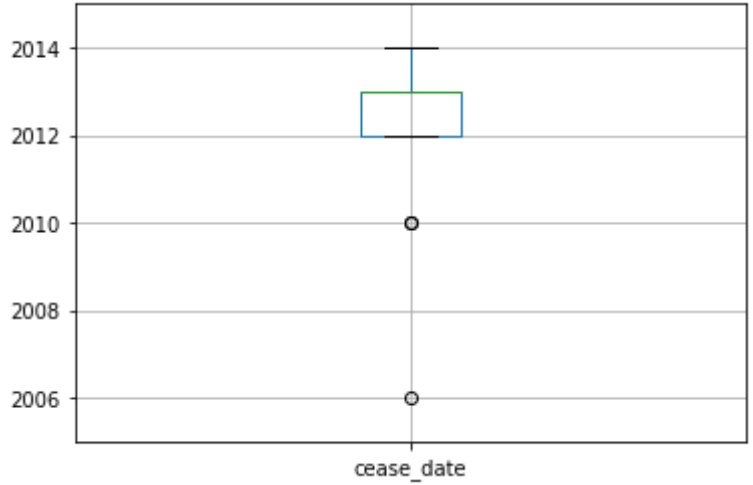
Out[23]:
```
2009.0      2
2013.0     55
2010.0     68
2012.0     94
2011.0    116
Name: cease_date, dtype: int64
```

In [24]:
```
import matplotlib.pyplot as plt
pv_dete_resignations = dete_resignations.pivot_table(values= 'cease_dat
e', index='dete_start_date')
pv_dete_resignations.plot(kind='box')
plt.show()
```



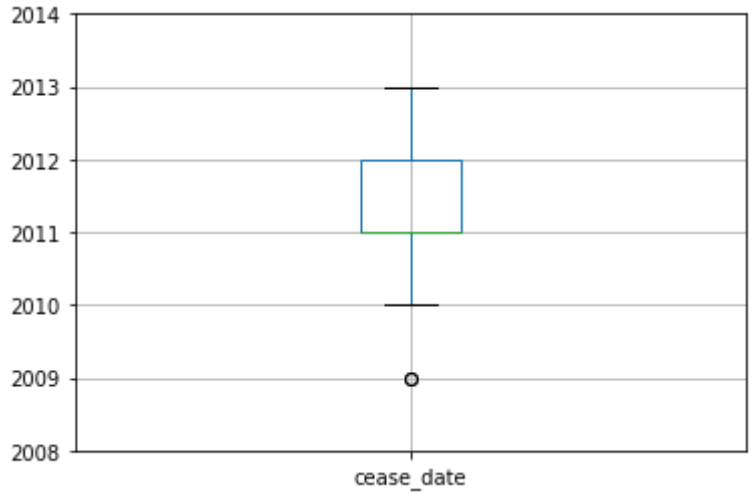
In [25]:
```
dete_resignations.boxplot(['cease_date']).set_ylim(2005,2015)
```

Out[25]: (2005, 2015)



In [26]:
```
tafe_resignations.boxplot('cease_date').set_ylim(2008,2014)
```

Out[26]:
```
(2008, 2014)
```



Now that we've verified the years in the dete_resignations dataframe, we'll use them to create a new column. Recall that our end goal is to answer the following question: . Are employees who have only worked for the institutes for a short period of time resigning due to some kind of dissatisfaction? What about employees who have been at the job longer? so the length of time an employee spent in a workplace is referred to as their years of service which already exist in tafe_resignations contains a "service" column as we renamed it to institute_service. In other word, we 'll create a corresponding institute_service column in dete_resignations as didn't exists.

Greating institute_service column through Subtracing the dete_start_date from the cease_date. Assign the result to a new column named institute_service.

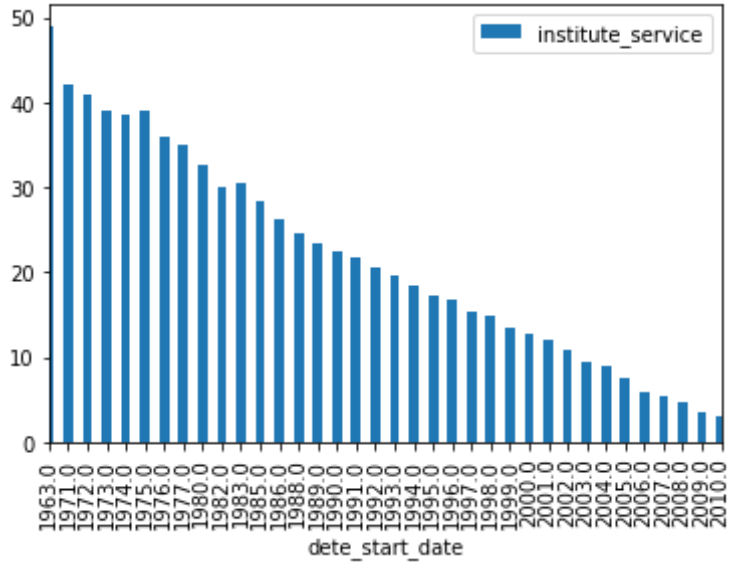
In [27]:
```
# Calculate the length of time an employee spent in their respective wo
rkplace and create a new column.
dete_resignations['institute_service'] = dete_resignations['cease_date'
] - dete_resignations['dete_start_date']

# Quick check of the result.
dete_resignations['institute_service'].head()
```

Out[27]:
```
3      7.0
5     18.0
8      3.0
9     15.0
11     3.0
Name: institute_service, dtype: float64
```

In [28]:
```
pv_dete_resignations = dete_resignations.pivot_table(values='institute_
service',index='dete_start_date', aggfunc='mean')
pv_dete_resignations.plot(kind='bar').set_xlim(0,35)
```

Out[28]:
```
(0, 35)
```



Distinguishing dissatisfied employees.

Next, we'll identify any employees who resigned because they were dissatisfied. Below are the columns we'll use to categorize employees as "dissatisfied" from each dataframe:

```
    tafe_survey_updated:
        Contributing Factors. Dissatisfaction
        Contributing Factors. Job Dissatisfaction
    dafe_survey_updated:
        job_dissatisfaction
        dissatisfaction_with_the_department
        physical_work_environment
        lack_of_recognition
        lack_of_job_security
        work_location
        employment_conditions
        work_life_balance
        workload
```

If the employee indicated any of the factors above caused them to resign, we'll mark them as dissatisfied in a new column. After our changes, the new dissatisfied column will contain just the following values:

```
    True: indicates a person resigned because they were dissatisfied
     in some way
    False: indicates a person resigned because of a reason other tha
    n dissatisfaction with the job
    NaN: indicates the value is missing
```

In [29]:
```
# Check the unique values

tafe_resignations['Contributing Factors. Dissatisfaction'].value_counts
()
```

Out[29]:
```
-                                     277
Contributing Factors. Dissatisfaction      55
Name: Contributing Factors. Dissatisfaction, dtype: int64
```

In [30]:
```
# Check the unique values

tafe_resignations['Contributing Factors. Job Dissatisfaction'].value_co
unts()
```

Out[30]:
```
-                      270
Job Dissatisfaction     62
Name: Contributing Factors. Job Dissatisfaction, dtype: int64
```

In [31]:
```
# Update the values in the contributing factors columns to be either Tr
ue, False, or NaN.

def update_vals(x):

    if x=='-':
        return False
    elif pd.isnull(x):
        return np.nan
    else:
        return True


tafe_resignations['dissatisfied'] = tafe_resignations[['Contributing Fa
ctors. Dissatisfaction', 'Contributing Factors. Job Dissatisfaction']].
applymap(update_vals).any(1, skipna=False)

tafe_resignations_up = tafe_resignations.copy()

# any(1, skipna=False) to include empty value as True.
```

In [32]:
```
tafe_resignations_update = tafe_resignations.copy()

# Check the unique values after the updates
tafe_resignations_up['dissatisfied'].value_counts(dropna=False)
```

Out[32]:
```
False    241
True      91
True       8
Name: dissatisfied, dtype: int64
```

In [33]:
```
# Update the values in columns related to dissatisfaction to be either
 True, False, or NaN
dissatisfied_reasons = ['job_dissatisfaction',
       'dissatisfaction_with_the_department', 'physical_work_environmen
t',
       'lack_of_recognition', 'lack_of_job_security', 'work_location',
       'employment_conditions', 'work_life_balance',
       'workload']
dete_resignations['dissatisfied'] = dete_resignations[dissatisfied_reas
ons].any(axis=1, skipna=False)

dete_resignations_up = dete_resignations.copy()

dete_resignations_up['dissatisfied'].value_counts()
```

Out[33]:
```
False    162
True     149
Name: dissatisfied, dtype: int64
```

Combining Data.

Below, we'll add an institute column so that we can differentiate the data from each survey after we combine them. Then, we'll combine the dataframes and drop any remaining columns we don't need.

In [34]:
```
dete_resignations_up['institute'] = 'DETE'
tafe_resignations_up['institute'] = 'TAFE'
```

In [37]:
```
# Combine the dataframes.
combined = pd.concat([tafe_resignations_up,dete_resignations_up], ignor
e_index=True)
```

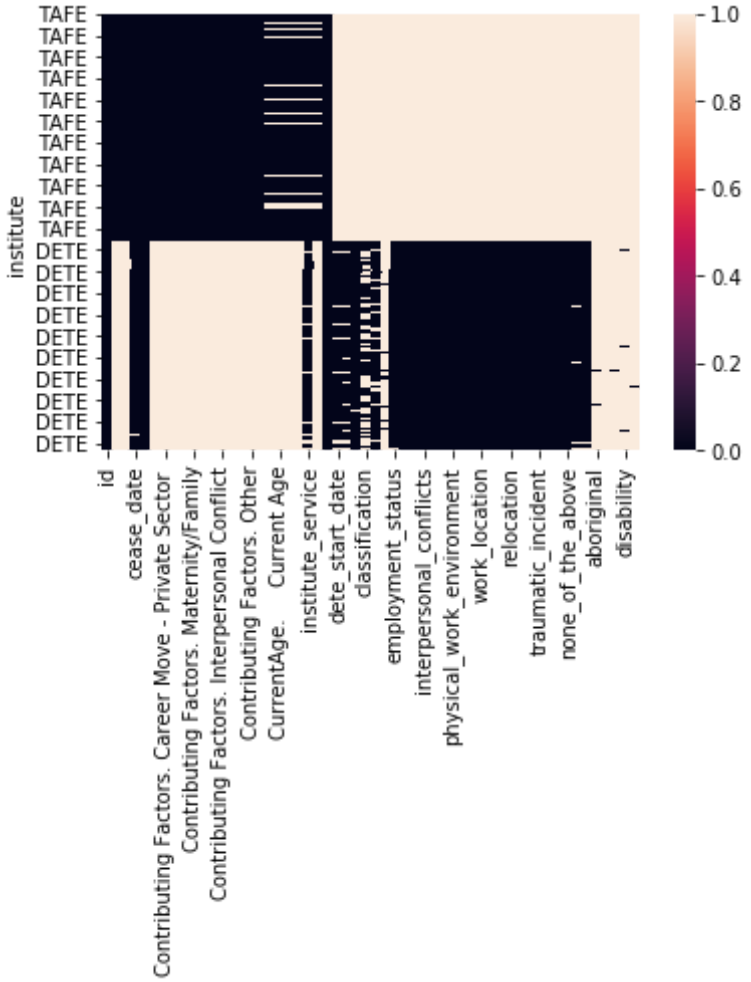
In [38]:
```
#  Find out the number of non null values of each cloumns.
combined.notnull().sum().sort_values()
```

Out[38]:
```
torres_strait                                          0
south_sea                                              3
aboriginal                                             7
disability                                             8
nesb                                                   9
business_unit                                         32
classification                                       161
region                                               265
role_start_date                                      271
dete_start_date                                      283
Gender.     What is your Gender?                     290
CurrentAge.     Current Age                          290
Employment Type.     Employment Type                 290
Classification.     Classification                   290
role_service                                         290
gender                                               302
age                                                  306
employment_status                                    307
position                                             308
interpersonal_conflicts                              311
career_move_to_private_sector                        311
career_move_to_public_sector                         311
physical_work_environment                            311
lack_of_recognition                                  311
lack_of_job_security                                 311
maternity/family                                     311
employment_conditions                                311
job_dissatisfaction                                  311
relocation                                           311
study/travel                                         311
ill_health                                           311
traumatic_incident                                   311
work_life_balance                                    311
workload                                             311
none_of_the_above                                    311
work_location                                        311
dissatisfaction_with_the_department                  311
Contributing Factors. Career Move - Self-employment  332
Contributing Factors. Interpersonal Conflict         332
```

```
Contributing Factors. Career Move - Public Sector      332
Contributing Factors. NONE                             332
Contributing Factors. Other                            332
Contributing Factors. Travel                           332
Contributing Factors. Study                            332
Contributing Factors. Career Move - Private Sector     332
Contributing Factors. Job Dissatisfaction              332
Contributing Factors. Dissatisfaction                  332
Contributing Factors. Maternity/Family                 332
Contributing Factors. Ill Health                       332
Institute                                              340
WorkArea                                               340
institute_service                                      563
cease_date                                             635
dissatisfied                                           643
separationtype                                         651
institute                                              651
id                                                     651
dtype: int64
```

In [39]:
```
# versualizing missing values.
com = combined.set_index('institute')
import seaborn as sns
sns.heatmap(com.isnull())
```

Out[39]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fda08431250>
```



In [40]:
```
# Let's drop the columns which including less than 500 non null values.
combined_updated = combined.dropna(thresh=500, axis=1).copy()
combined_updated
```

Out[40]:

	id	cease_date	separationtype	institute_service	dissatisfied	institute
0	634139903350000000	2010.0	Resignation	NaN	False	TAFE
1	634146578511788000	2010.0	Resignation	3-4	False	TAFE

	id	cease_date	separationtype	institute_service	dissatisfied	institute
2	634147506906311000	2010.0	Resignation	7-10	False	TAFE
3	634152007975694000	2010.0	Resignation	3-4	False	TAFE
4	634153745310374000	2010.0	Resignation	3-4	False	TAFE
...	...	...	...	...	...	...
646	810	2013.0	Resignation	3	False	DETE
647	817	2014.0	Resignation	2	False	DETE
648	818	2014.0	Resignation	2	False	DETE
649	821	2014.0	Resignation	5	True	DETE
650	823	2013.0	Resignation	NaN	False	DETE

651 rows × 6 columns

Cleaning the institution_service column.

We'll use the slightly modified definitions below:

```
New: Less than 3 years at a company
Experienced: 3-6 years at a company
Established: 7-10 years at a company
Veteran: 11 or more years at a company
```

Let's categorize the values in the institute_service column using the definitions above.

In [41]:
```
# Check the unique values.
combined_updated['institute_service'].value_counts()
```

Out[41]:
```
Less than 1 year      73
1-2                   64
3-4                   63
```

5-6	33
11-20	26
5.0	23
1.0	22
7-10	21
0.0	20
3.0	20
6.0	17
4.0	16
9.0	14
2.0	14
7.0	13
More than 20 years	10
8.0	8
13.0	8
15.0	7
20.0	7
10.0	6
12.0	6
14.0	6
17.0	6
22.0	6
16.0	5
18.0	5
11.0	4
24.0	4
23.0	4
19.0	3
39.0	3
21.0	3
32.0	3
28.0	2
30.0	2
26.0	2
36.0	2
25.0	2
27.0	1
29.0	1
31.0	1

```
33.0                    1
34.0                    1
35.0                    1
38.0                    1
41.0                    1
42.0                    1
49.0                    1
Name: institute_service, dtype: int64
```

In [42]:
```
# Extract the years from institute_service and convert the type to float.
combined_updated['institute_service_up'] = combined_updated['institute_service'].astype('str').str.extract(r'(\d+)')
combined_updated['institute_service_up'] = combined_updated['institute_service_up'].astype('float')

# Check the years extracted are correct.
combined_updated['institute_service_up'].value_counts()
```

Out[42]:
```
1.0     159
3.0      83
5.0      56
7.0      34
11.0     30
0.0      20
6.0      17
20.0     17
4.0      16
9.0      14
2.0      14
8.0       8
13.0      8
15.0      7
22.0      6
14.0      6
17.0      6
12.0      6
10.0      6
18.0      5
16.0      5
```

```
23.0      4
24.0      4
21.0      3
19.0      3
32.0      3
39.0      3
30.0      2
26.0      2
36.0      2
28.0      2
25.0      2
35.0      1
38.0      1
34.0      1
33.0      1
49.0      1
41.0      1
27.0      1
42.0      1
29.0      1
31.0      1
Name: institute_service_up, dtype: int64
```

In [43]:

```
def extract_var(element):
    if pd.isnull(element):
        return np.nan
    elif element <= 3:
        return 'New'
    elif 3< element <=6:
        return 'Experienced'
    elif 7 <= element <=10:
        return "Established"
    else:
        return 'Veteran'

combined_updated['institute_cat'] = combined_updated['institute_service
_up'].map(extract_var)
combined_updated['institute_cat'].value_counts()
```

Out[43]:
```
New            276
Veteran        136
Experienced     89
Established     62
Name: institute_cat, dtype: int64
```

Perform Some Initial Analysis

Finally, we'll replace the missing values in the dissatisfied column with the most frequent value, False. Then, we'll calculate the percentage of employees who resigned due to dissatisfaction in each service_cat group and plot the results.

Note that since we still have additional missing values left to deal with, this is meant to be an initial introduction to the analysis, not the final analysis.

In [44]:
```
# Verify the unique values.
combined_updated['dissatisfied'].value_counts(dropna=False)
```

Out[44]:
```
False    403
True     240
True       8
Name: dissatisfied, dtype: int64
```

In [45]:
```
# Replace missing values with the most frequent value, False
combined_updated['dissatisfied'] = combined_updated['dissatisfied'].fil
lna(False)
combined_updated['dissatisfied'].value_counts(dropna=False)
```

Out[45]:
```
False    411
True     240
Name: dissatisfied, dtype: int64
```

In [46]:
```
# Calculate the percentage of employees who resigned due to dissatisfac
tion in each category
```

```
pv_combined_update = combined_updated.pivot_table(index='institute_cat'
, values = 'dissatisfied',aggfunc=np.mean)
pv_combined_update
```

Out[46]:

institute_cat	dissatisfied
Established	0.516129
Experienced	0.404494
New	0.289855
Veteran	0.485294

In [47]:
```
# Visualizing the results.
%matplotlib inline
pv_combined_update.plot(kind='bar',rot=45)
```

Out[47]: <matplotlib.axes._subplots.AxesSubplot at 0x7fda08667950>