In [1]:
```
import pandas as pd

df_3_97 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/1997/Form3.csv', encoding = "ISO-885
df_3_98 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/1998/Form3.csv', encoding = "ISO-885
df_3_99 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/1999/Form3.csv', encoding = "ISO-885
df_3_00 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2000/Form3.csv', encoding = "ISO-885
df_3_01 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2001/Form3.csv', encoding = "ISO-885
df_3_02 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2002/Form3.csv', encoding = "ISO-885
df_3_03 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2003/Form3.csv', encoding = "ISO-885
df_3_04 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2004/Form3.csv', encoding = "ISO-885
df_3_05 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2005/Form3.csv', encoding = "ISO-885
df_3_06 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2006/Form3.csv', encoding = "ISO-885
df_3_07 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2007/Form3.csv', encoding = "ISO-885
df_3_08 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2008/Form3.csv', encoding = "ISO-885
df_3_09 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2009/Form3.csv', encoding = "ISO-885
df_3_10 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2010/Form3.csv', encoding = "ISO-885
df_3_11 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2011/Form3.csv', encoding = "ISO-885
df_3_12 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2012/Form3.csv', encoding = "ISO-885
df_3_13 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2013/Form3.csv', encoding = "ISO-885
df_3_14 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2014/Form3.csv', encoding = "ISO-885
df_3_15 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2015/Form3.csv', encoding = "ISO-885
df_3_16 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2016/Form3.csv', encoding = "ISO-885
df_3_17 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2017/Form3.csv', encoding = "ISO-885
df_3_18 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2018/Form3.csv', encoding = "ISO-885
df_3_19 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2019/Form3.csv', encoding = "ISO-885
```

In [2]:
```
import matplotlib.colors as mplc
import matplotlib.patches as patches
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
import statsmodels.formula.api as sm
```

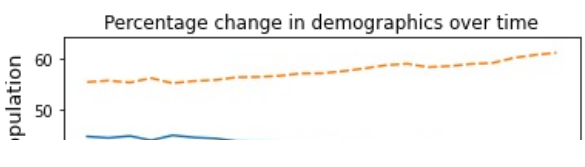
In [6]:
```
df397 = df_3_97.loc[df_3_97["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df398 = df_3_98.loc[df_3_98["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df399 = df_3_99.loc[df_3_99["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df300 = df_3_00.loc[df_3_00["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df301 = df_3_01.loc[df_3_01["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df302 = df_3_02.loc[df_3_02["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df303 = df_3_03.loc[df_3_03["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df304 = df_3_04.loc[df_3_04["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df305 = df_3_05.loc[df_3_05["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df306 = df_3_06.loc[df_3_06["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df307 = df_3_07.loc[df_3_07["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df308 = df_3_08.loc[df_3_08["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df309 = df_3_09.loc[df_3_09["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df310 = df_3_10.loc[df_3_10["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df311 = df_3_11.loc[df_3_11["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df312 = df_3_12.loc[df_3_12["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df313 = df_3_13.loc[df_3_13["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df314 = df_3_14.loc[df_3_14["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df315 = df_3_15.loc[df_3_15["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df316 = df_3_16.loc[df_3_16["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df317 = df_3_17.loc[df_3_17["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df318 = df_3_18.loc[df_3_18["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df319 = df_3_19.loc[df_3_19["SALARYRANGE($)"] ==  "Overall"].groupby("CALENDARYEAR").sum()
```

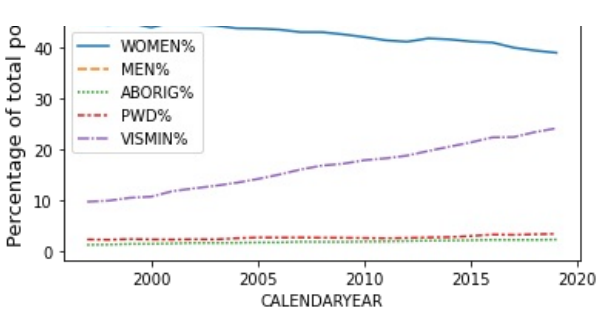
In [21]:
```
df3final = pd.concat([df397,df398,df399,df300,df301,df302,df303,df304,df305,df306,df307,df308,df309,df310,
          df311,df312,df313,df314,df315,df316,df317,df318,df319])
df3final["MEN%"] = df3final["ALLMENCOUNT"]/df3final["ALLCOUNT"] *100
df3final["WOMEN%"]= df3final["ALLWOMENCOUNT"]/df3final["ALLCOUNT"] *100
df3final["ABORIG%"]= df3final["ABORIGALLCOUNT"]/df3final["ALLCOUNT"] *100
df3final["PWD%"] = df3final["PWDALLCOUNT"]/df3final["ALLCOUNT"] *100
df3final["VISMIN%"] = df3final["VISMINALLCOUNT"]/df3final["ALLCOUNT"] *100
```

In [25]:
```
df3finalfinal = df3final[["WOMEN%", "MEN%", "ABORIG%", "PWD%", "VISMIN%"]]
df3final3 = df3final[["ABORIG%", "PWD%"]]
```

In [26]:
```
df3plot = sns.lineplot(data = df3finalfinal)
df3plot.set_ylabel("Percentage of total population", fontsize = 13)
df3plot.set_title("Percentage change in demographics over time")
```

Out[26]:
```
Text(0.5, 1.0, 'Percentage change in demographics over time')
```



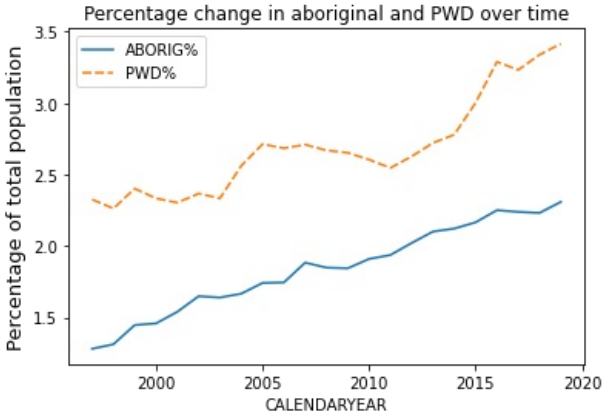


In [27]:

```
df3plot3 = sns.lineplot(data = df3final3)
df3plot3.set_ylabel("Percentage of total population", fontsize = 13)
df3plot3.set_title("Percentage change in aboriginal and PWD over time")
```

Out[27]:

```
Text(0.5, 1.0, 'Percentage change in aboriginal and PWD over time')
```



In [ ]: