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

df_5_97 = pd.read_csv('C:\\Users\\amber\\Documents\\R\\CANDEV\\Form5.csv', encoding = "ISO-8859-1")
df_5_98 = pd.read_csv('C:\\Users\\amber\\Documents\\R\\CANDEV\\Form5 98.csv', encoding = "ISO-8859-1")
df_5_99 = pd.read_csv('C:\\Users\\amber\\Documents\\R\\CANDEV\\Form5 99.csv', encoding = "ISO-8859-1")
df_5_00 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2000/Form5.csv', encoding = "ISO-885
df_5_01 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2001/Form5.csv', encoding = "ISO-885
df_5_02 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2002/Form5.csv', encoding = "ISO-885
df_5_03 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2003/Form5.csv', encoding = "ISO-885
df_5_04 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2004/Form5.csv', encoding = "ISO-885
df_5_05 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2005/Form5.csv', encoding = "ISO-885
df_5_06 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2006/Form5.csv', encoding = "ISO-885
df_5_07 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2007/Form5.csv', encoding = "ISO-885
df_5_08 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2008/Form5.csv', encoding = "ISO-885
df_5_09 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2009/Form5.csv', encoding = "ISO-885
df_5_10 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2010/Form5.csv', encoding = "ISO-885
df_5_11 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2011/Form5.csv', encoding = "ISO-885
df_5_12 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2012/Form5.csv', encoding = "ISO-885
df_5_13 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2013/Form5.csv', encoding = "ISO-885
df_5_14 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2014/Form5.csv', encoding = "ISO-885
df_5_15 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2015/Form5.csv', encoding = "ISO-885
df_5_16 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2016/Form5.csv', encoding = "ISO-885
df_5_17 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2017/Form5.csv', encoding = "ISO-885
df_5_18 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2018/Form5.csv', encoding = "ISO-885
df_5_19 = pd.read_csv('http://www.esdc.gc.ca/ouvert-open/labour-travail/leep/2019/Form5.csv', encoding = "ISO-885
```

In [12]:
```
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
```

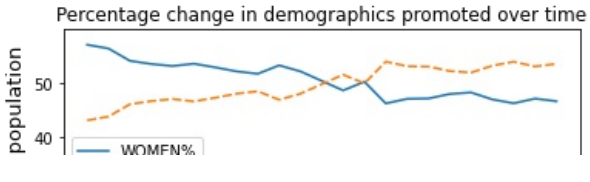
In [75]:
```
df597 = df_5_97.loc[df_5_97["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df598 = df_5_98.loc[df_5_98["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df599 = df_5_99.loc[df_5_99["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df500 = df_5_00.loc[df_5_00["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df501 = df_5_01.loc[df_5_01["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df502 = df_5_02.loc[df_5_02["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df503 = df_5_03.loc[df_5_03["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df504 = df_5_04.loc[df_5_04["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df505 = df_5_05.loc[df_5_05["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df506 = df_5_06.loc[df_5_06["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df507 = df_5_07.loc[df_5_07["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df508 = df_5_08.loc[df_5_08["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df509 = df_5_09.loc[df_5_09["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df510 = df_5_10.loc[df_5_10["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df511 = df_5_11.loc[df_5_11["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df512 = df_5_12.loc[df_5_12["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df513 = df_5_13.loc[df_5_13["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df514 = df_5_14.loc[df_5_14["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df515 = df_5_15.loc[df_5_15["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df516 = df_5_16.loc[df_5_16["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df517 = df_5_17.loc[df_5_17["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df518 = df_5_18.loc[df_5_18["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
df519 = df_5_19.loc[df_5_19["OCCGROUP"] ==  "Overall"].groupby("CALENDARYEAR").sum()
```

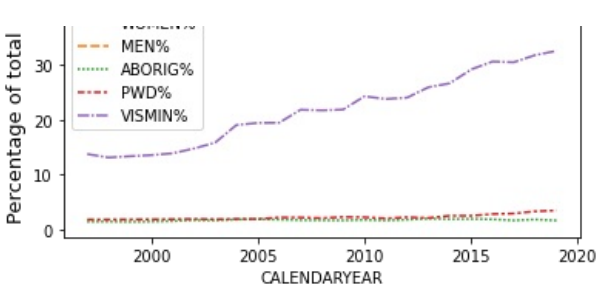
In [78]:
```
df5final = pd.concat([df597,df598,df599,df500,df501,df502,df503,df504,df505,df506,df507,df508,df509,df510,
          df511,df512,df513,df514,df515,df516,df517,df518,df519])
df5final["MEN%"] = df5final["TOTALMENPROMOTIONS"]/df5final["TOTALPROMOTIONS"] *100
df5final["WOMEN%"]= df5final["TOTALWOMENPROMOTIONS"]/df5final["TOTALPROMOTIONS"] *100
df5final["ABORIG%"]= df5final["TOTALABORIGALLPROMOTIONS"]/df5final["TOTALPROMOTIONS"] *100
df5final["PWD%"] = df5final["TOTALPWDALLPROMOTIONS"]/df5final["TOTALPROMOTIONS"] *100
df5final["VISMIN%"] = df5final["VISMINALLCOUNT"]/df5final["TOTALPROMOTIONS"] *100
```

In [81]:
```
df5finalfinal = df5final[["WOMEN%", "MEN%", "ABORIG%", "PWD%", "VISMIN%"]]
df5final3 = df5final[["ABORIG%", "PWD%"]]
```

In [82]:
```
df5plot = sns.lineplot(data = df5finalfinal)
df5plot.set_ylabel("Percentage of total population", fontsize = 13)
df5plot.set_title("Percentage change in demographics promoted over time")
```

Out[82]:
```
Text(0.5, 1.0, 'Percentage change in demographics promoted over time')
```



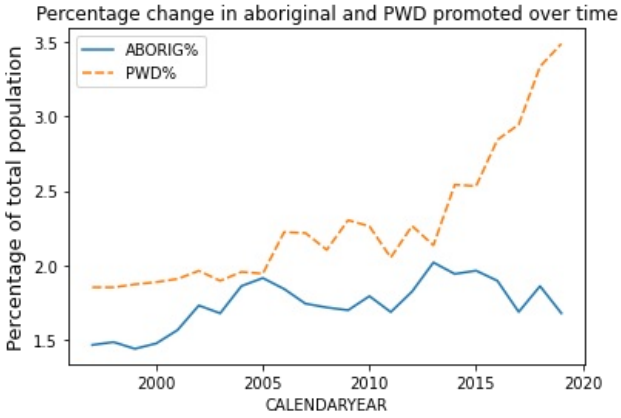


In [83]:

```
df5plot3 = sns.lineplot(data = df5final3)
df5plot3.set_ylabel("Percentage of total population", fontsize = 13)
df5plot3.set_title("Percentage change in aboriginal and PWD promoted over time")
```

Out[83]:

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



In [ ]: