**Option 1**

I would look at [the pandas library](http://pandas.pydata.org/). For example, if we start from two similar files:

>>> import pandas as pd

>>>

>>> df0 = pd.ExcelFile("id\_data1.xls").parse("Sheet1")

>>> df1 = pd.ExcelFile("id\_data2.xls").parse("Sheet1")

>>> df0 = df0.set\_index("ID")

>>> df1 = df1.set\_index("ID")

>>> df0

A B

ID

1 a e

2 b f

3 c g

4 d h

>>> df1

A B

ID

1 a e

2 b nolongerf

4 d h

5 g h

We can align them, look at the differences, and save the result to an excel file:

>>> a0, a1 = df0.align(df1)

>>> different = (a0 != a1).any(axis=1)

>>> comp = a0[different].join(a1[different], lsuffix='\_old', rsuffix='\_new')

>>> comp

A\_old B\_old A\_new B\_new

ID

2 b f b nolongerf

3 c g NaN NaN

5 NaN NaN g h

>>> comp.to\_excel("comparison.xls")

>>>

**Option 2**

from pandas import read\_excel

import numpy as np

df = read\_excel('excel\_data.xlsx', names=['A','B'], header=None)

df

#only in list A

np.setdiff1d(df['A'], df['B'])

#only in list B

np.setdiff1d(df['B'], df['A'])

Option 3

from itertools import izip\_longest

import xlrd

rb1 = xlrd.open\_workbook('file1.xlsx')

rb2 = xlrd.open\_workbook('file2.xlsx')

sheet1 = rb1.sheet\_by\_index(0)

sheet2 = rb2.sheet\_by\_index(0)

for rownum in range(max(sheet1.nrows, sheet2.nrows)):

if rownum < sheet1.nrows:

row\_rb1 = sheet1.row\_values(rownum)

row\_rb2 = sheet2.row\_values(rownum)

for colnum, (c1, c2) in enumerate(izip\_longest(row\_rb1, row\_rb2)):

if c1 != c2:

print "Row {} Col {} - {} != {}".format(rownum+1, colnum+1, c1, c2)

else:

print "Row {} missing".format(rownum+1)