

```
In [58]: F = len('Katarzyna')
L = len('Watorska')
N = (F+L)*100
```

Exercise 1

1.Compile the code in code1.stan and code2.stan

```
In [59]: from cmdstanpy import CmdStanModel
import arviz as az
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
In [60]: model1 = CmdStanModel(stan_file='code1.stan')
```

INFO:cmdstanpy:found newer exe file, not recompiling

```
In [61]: model2 = CmdStanModel(stan_file='code2.stan')
```

INFO:cmdstanpy:found newer exe file, not recompiling

2.Generate data for the rest of the exercises.

```
In [62]: stan_data = {'N':N}
```

```
In [63]: samples_model1 = model1.sample(data=stan_data)
samples_model2 = model2.sample(data=stan_data)
```

```
y1 = samples_model1.stan_variable(var='y')
y2 = samples_model2.stan_variable(var='y')
```

INFO:cmdstanpy:CmdStan start processing

chain 1		00:00	Status
chain 2		00:00	Status
chain 3		00:00	Status
chain 4		00:00	Status

INFO:cmdstanpy:CmdStan done processing.

INFO:cmdstanpy:CmdStan start processing

chain 1		00:00	Status
chain 2		00:00	Status
chain 3		00:00	Status
chain 4		00:00	Status

INFO:cmdstanpy:CmdStan done processing.

Exercise 2

```
In [ ]: 1.Compile both models
```

```
In [3]: model3 = CmdStanModel(stan_file='code3.stan')
model4 = CmdStanModel(stan_file='code4.stan')
```

```
INFO:cmdstanpy:compiling stan file /home/kasia/Documents/DataAnalytics/La
b7/code3.stan to exe file /home/kasia/Documents/DataAnalytics/Lab7/code3
INFO:cmdstanpy:compiled model executable: /home/kasia/Documents/DataAnaly
tics/Lab7/code3
INFO:cmdstanpy:compiling stan file /home/kasia/Documents/DataAnalytics/La
b7/code4.stan to exe file /home/kasia/Documents/DataAnalytics/Lab7/code4
INFO:cmdstanpy:compiled model executable: /home/kasia/Documents/DataAnaly
tics/Lab7/code4
```

```
In [40]: stan_data3={'N': N, 'y': y1[0]}

fit3 = model3.sample(data=stan_data3)
```

```
INFO:cmdstanpy:CmdStan start processing
chain 1 |           | 00:00 Status
chain 2 |           | 00:00 Status
chain 3 |           | 00:00 Status
chain 4 |           | 00:00 Status
```

```
INFO:cmdstanpy:CmdStan done processing.
```

```
In [18]: stan_data4={'N': N, 'y': y2[0]}

fit4 = model4.sample(data=stan_data4)
```

```
INFO:cmdstanpy:CmdStan start processing
chain 1 |           | 00:00 Status
chain 2 |           | 00:00 Status
chain 3 |           | 00:00 Status
chain 4 |           | 00:00 Status
```

```
INFO:cmdstanpy:CmdStan done processing.
```

3.Compare both models

```
In [64]: compare_dict = {"model3" : fit3, "model4" : fit4}
model34_loo=az.compare(compare_dict, ic='loo')
az.plot_compare(model34_loo)
```

```
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:694:
UserWarning: Estimated shape parameter of Pareto distribution is greater
than 0.7 for one or more samples. You should consider using a more robust
model, this is because importance sampling is less likely to work well if
the marginal posterior and LOO posterior are very different. This is more
likely to happen with a non-robust model and highly influential observati
ons.
```

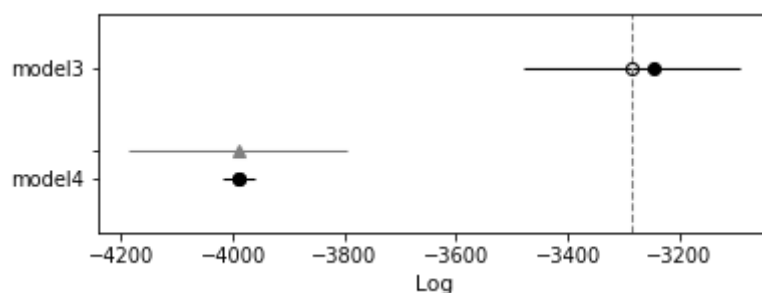
```
warnings.warn(
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
```

```
ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_nam
e=var_name)])
```

```
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
```

```
ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_nam
e=var_name)])
```

```
Out[64]: <AxesSubplot:xlabel='Log'>
```



```
In [65]: model34_waic = az.compare(compare_dict, ic='waic')
         az.plot_compare(model34_waic)
```

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:1458:
UserWarning: For one or more samples the posterior variance of the log predictive densities exceeds 0.4. This could be indication of WAIC starting to fail.

See <http://arxiv.org/abs/1507.04544> for details

warnings.warn(

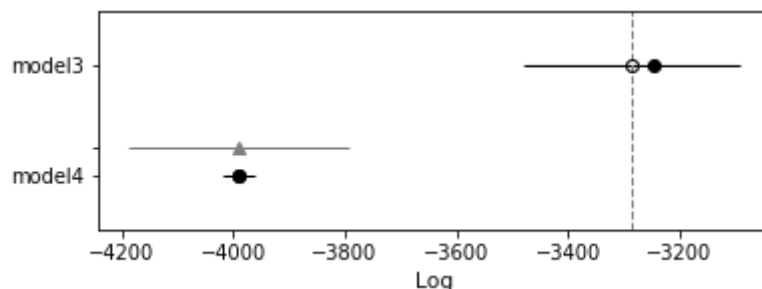
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])

```
Out[65]: <AxesSubplot:xlabel='Log'>
```



```
In [23]: model34 = az.compare(compare_dict)
```

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:694:
UserWarning: Estimated shape parameter of Pareto distribution is greater than 0.7 for one or more samples. You should consider using a more robust model, this is because importance sampling is less likely to work well if the marginal posterior and L00 posterior are very different. This is more likely to happen with a non-robust model and highly influential observations.

warnings.warn(

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])

```
Out [23]:
```

	rank	loo	p_loo	d_loo	weight	se	dse	warning
model3	0	-3109.012067	34.379023	0.000000	0.986629	175.879744	0.000000	True
model4	1	-3990.097924	1.930443	881.085857	0.013371	27.959649	178.12944	False

Exercise 3

1. Compile the model

```
In [32]: model5 = CmdStanModel(stan_file='code5.stan')
```

```
INFO:cmdstanpy:compiling stan file /home/kasia/Documents/DataAnalytics/La
b7/code5.stan to exe file /home/kasia/Documents/DataAnalytics/Lab7/code5
INFO:cmdstanpy:compiled model executable: /home/kasia/Documents/DataAnaly
tics/Lab7/code5
```

2. Compare the model for 1, 2, 3 predictors.

```
In [57]: y_hat=fit4.stan_variable('y_hat')
X1=y_hat[0:1].transpose()
data5_1= {'N':N, 'K':1, 'y':y2[0] , 'X':X1 }
fit5_1=model5.sample(data=data5_1)
X2=y_hat[0:2].transpose()
data5_2= {'N':N, 'K':2, 'y':y2[0] , 'X':X2 }
fit5_2=model5.sample(data=data5_2)
X3=y_hat[0:3].transpose()
data5_3= {'N':N, 'K':3, 'y':y2[0] , 'X':X3 }
fit5_3=model5.sample(data=data5_3)

compare_dict2 = {"model5_1" : fit5_1, "model5_2" : fit5_2, "model5_3": fi
model5_loo=az.compare(compare_dict2, ic='loo')
az.plot_compare(model5_loo)
```

```
INFO:cmdstanpy:CmdStan start processing
chain 1 | | 00:00 Status
chain 2 | | 00:00 Status
chain 3 | | 00:00 Status
chain 4 | | 00:00 Status
```

```
INFO:cmdstanpy:CmdStan done processing.
```

```
INFO:cmdstanpy:CmdStan start processing
chain 1 | | 00:00 Status
chain 2 | | 00:00 Status
chain 3 | | 00:00 Status
chain 4 | | 00:00 Status
```

```
INFO:cmdstanpy:CmdStan done processing.
```

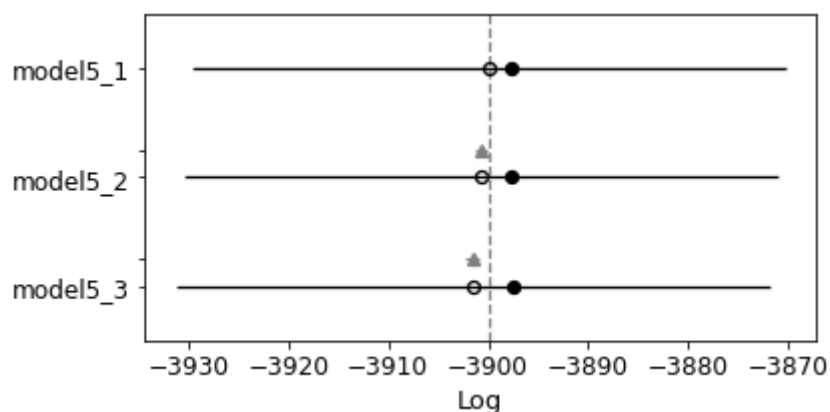
```
INFO:cmdstanpy:CmdStan start processing
chain 1 | | 00:00 Status
chain 2 | | 00:00 Status
chain 3 | | 00:00 Status
chain 4 | | 00:00 Status
```

```
INFO:cmdstanpy:CmdStan done processing.
```

```

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
    ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
    ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
    ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])
Out[57]: <AxesSubplot:xlabel='Log'>

```



```

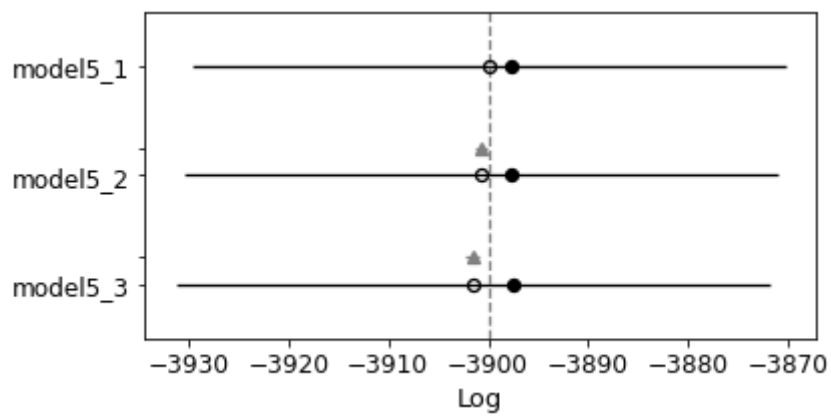
In [67]: model5_waic=az.compare(compare_dict2, ic='waic')
         az.plot_compare(model5_waic)

```

```

/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
    ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
    ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])
/home/kasia/.local/lib/python3.8/site-packages/arviz/stats/stats.py:248:
FutureWarning: The frame.append method is deprecated and will be removed
from pandas in a future version. Use pandas.concat instead.
    ics = ics.append([ic_func(dataset, pointwise=True, scale=scale, var_name=var_name)])
Out[67]: <AxesSubplot:xlabel='Log'>

```



In []: