Legend:
- on-off AWAC_GAE, DA False
- on-off AWAC_Q_lambda_Haru, DA False
- on-off AWAC_Q_lambda_Peng, DA False
- on-off AWAC_Q_lambda_TB, DA False
- PPO
- average reward data set