MC_2

```
library(MASS)
library(pROC)
library(dplyr)
library(glmnet)
library(tibble)
library(ggplot2)
```

Main Findings

1. When EPV is small (EPV=5), Ridge performs better overall than MLE, showing:

- higher or similar AUC
- lower Brier scores
- much lower variance
- more realistic estimates when predictors are correlated

2. At EPV=10, both methods perform comparably,

- Although, Ridge remains slightly more stable, especially in the presence of noisy predictors.

3. At EPV=50, MLE and Ridge are almost identical, because shrinkage is no longer needed.

4. Ridge's adaptive penalty (λ) adjusts appropriately:

- strong shrinkage at low EPV
- minimal shrinkage at high EPV

→ Ridge regression is advantageous when EPV is low or when predictors contain noise or are correlated. When EPV is sufficiently high (≥ 50), both methods perform equivalently.

```
summary_df <- results_df %>%
  group_by(model, EPV, P, rho, beta_pattern) %>%
  summarise(
    mean_AUC = mean(AUC),
    sd_AUC = sd(AUC),
    mean_Brier = mean(Brier),
    sd_Brier = sd(Brier),
    mean_lambda = mean(lambda, na.rm = TRUE),
    .groups = "drop"
  )

knitr::kable(summary_df, digits = 3)
```

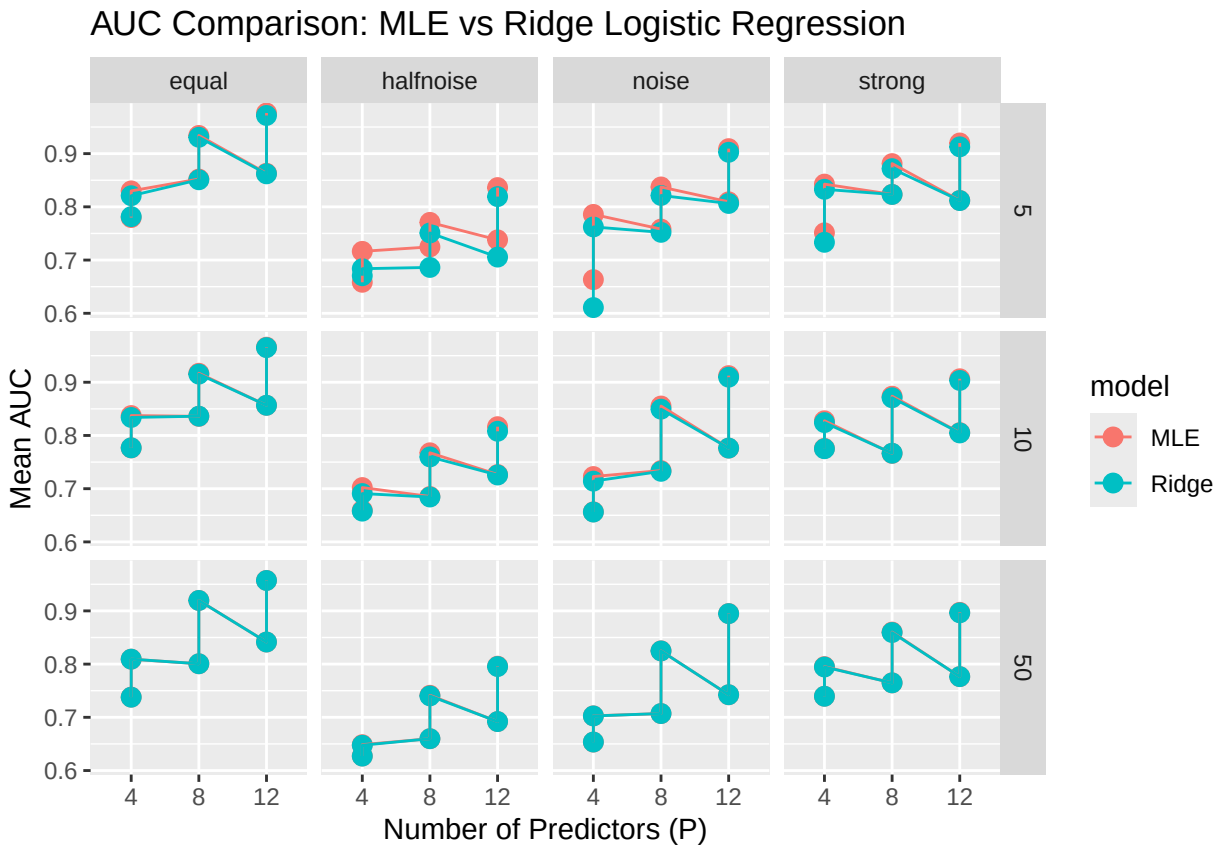
model	EPV	P	rho	beta_pattern	mean_AUC	sd_AUC	mean_Brier	sd_Brier	mean_lambda
MLE	5	4	0.0	equal	0.780	0.060	0.182	0.025	NaN
MLE	5	4	0.0	halfnoise	0.658	0.140	0.213	0.023	NaN
MLE	5	4	0.0	noise	0.663	0.053	0.223	0.018	NaN
MLE	5	4	0.0	strong	0.751	0.080	0.195	0.030	NaN
MLE	5	4	0.5	equal	0.830	0.071	0.160	0.036	NaN
MLE	5	4	0.5	halfnoise	0.716	0.076	0.200	0.028	NaN

model	EPV	P	rho	beta_pattern	mean_AUC	sd_AUC	mean_Brier	sd_Brier	mean_lambda
MLE	5	4	0.5	noise	0.786	0.070	0.184	0.033	NaN
MLE	5	4	0.5	strong	0.843	0.069	0.155	0.034	NaN
MLE	5	8	0.0	equal	0.852	0.036	0.155	0.021	NaN
MLE	5	8	0.0	halfnoise	0.725	0.045	0.208	0.017	NaN
MLE	5	8	0.0	noise	0.758	0.044	0.194	0.021	NaN
MLE	5	8	0.0	strong	0.823	0.035	0.168	0.021	NaN
MLE	5	8	0.5	equal	0.934	0.032	0.098	0.026	NaN
MLE	5	8	0.5	halfnoise	0.771	0.053	0.190	0.019	NaN
MLE	5	8	0.5	noise	0.837	0.044	0.159	0.024	NaN
MLE	5	8	0.5	strong	0.881	0.032	0.137	0.018	NaN
MLE	5	12	0.0	equal	0.863	0.028	0.148	0.015	NaN
MLE	5	12	0.0	halfnoise	0.738	0.053	0.205	0.017	NaN
MLE	5	12	0.0	noise	0.809	0.045	0.174	0.022	NaN
MLE	5	12	0.0	strong	0.812	0.049	0.174	0.022	NaN
MLE	5	12	0.5	equal	0.976	0.022	0.054	0.031	NaN
MLE	5	12	0.5	halfnoise	0.836	0.030	0.163	0.016	NaN
MLE	5	12	0.5	noise	0.909	0.025	0.121	0.017	NaN
MLE	5	12	0.5	strong	0.920	0.035	0.111	0.023	NaN
MLE	10	4	0.0	equal	0.777	0.044	0.190	0.017	NaN
MLE	10	4	0.0	halfnoise	0.659	0.040	0.228	0.008	NaN
MLE	10	4	0.0	noise	0.656	0.104	0.221	0.019	NaN
MLE	10	4	0.0	strong	0.776	0.036	0.186	0.015	NaN
MLE	10	4	0.5	equal	0.837	0.055	0.157	0.033	NaN
MLE	10	4	0.5	halfnoise	0.702	0.031	0.217	0.009	NaN
MLE	10	4	0.5	noise	0.723	0.043	0.204	0.015	NaN
MLE	10	4	0.5	strong	0.828	0.037	0.166	0.018	NaN
MLE	10	8	0.0	equal	0.836	0.028	0.165	0.014	NaN
MLE	10	8	0.0	halfnoise	0.685	0.047	0.221	0.014	NaN
MLE	10	8	0.0	noise	0.734	0.024	0.207	0.008	NaN
MLE	10	8	0.0	strong	0.766	0.023	0.195	0.009	NaN
MLE	10	8	0.5	equal	0.917	0.017	0.115	0.012	NaN
MLE	10	8	0.5	halfnoise	0.767	0.045	0.193	0.017	NaN
MLE	10	8	0.5	noise	0.855	0.038	0.152	0.022	NaN
MLE	10	8	0.5	strong	0.874	0.032	0.142	0.018	NaN
MLE	10	12	0.0	equal	0.857	0.027	0.153	0.015	NaN
MLE	10	12	0.0	halfnoise	0.727	0.018	0.211	0.006	NaN
MLE	10	12	0.0	noise	0.776	0.016	0.192	0.006	NaN
MLE	10	12	0.0	strong	0.805	0.032	0.179	0.013	NaN
MLE	10	12	0.5	equal	0.966	0.009	0.072	0.009	NaN
MLE	10	12	0.5	halfnoise	0.817	0.037	0.173	0.018	NaN
MLE	10	12	0.5	noise	0.912	0.025	0.117	0.019	NaN
MLE	10	12	0.5	strong	0.907	0.019	0.122	0.013	NaN
MLE	50	4	0.0	equal	0.738	0.021	0.207	0.008	NaN
MLE	50	4	0.0	halfnoise	0.627	0.029	0.237	0.005	NaN
MLE	50	4	0.0	noise	0.653	0.033	0.231	0.008	NaN
MLE	50	4	0.0	strong	0.740	0.023	0.206	0.008	NaN
MLE	50	4	0.5	equal	0.809	0.024	0.177	0.011	NaN
MLE	50	4	0.5	halfnoise	0.649	0.024	0.232	0.006	NaN
MLE	50	4	0.5	noise	0.703	0.028	0.217	0.008	NaN
MLE	50	4	0.5	strong	0.795	0.019	0.184	0.008	NaN
MLE	50	8	0.0	equal	0.801	0.010	0.182	0.005	NaN
MLE	50	8	0.0	halfnoise	0.660	0.014	0.230	0.003	NaN

model	EPV	P	rho	beta_pattern	mean_AUC	sd_AUC	mean_Brier	sd_Brier	mean_lambda
MLE	50	8	0.0	noise	0.707	0.014	0.217	0.004	NaN
MLE	50	8	0.0	strong	0.765	0.015	0.197	0.006	NaN
MLE	50	8	0.5	equal	0.920	0.012	0.113	0.009	NaN
MLE	50	8	0.5	halfnoise	0.741	0.010	0.206	0.003	NaN
MLE	50	8	0.5	noise	0.825	0.019	0.170	0.009	NaN
MLE	50	8	0.5	strong	0.860	0.014	0.152	0.008	NaN
MLE	50	12	0.0	equal	0.842	0.012	0.162	0.006	NaN
MLE	50	12	0.0	halfnoise	0.692	0.016	0.222	0.005	NaN
MLE	50	12	0.0	noise	0.742	0.010	0.206	0.004	NaN
MLE	50	12	0.0	strong	0.776	0.011	0.192	0.004	NaN
MLE	50	12	0.5	equal	0.957	0.007	0.082	0.008	NaN
MLE	50	12	0.5	halfnoise	0.796	0.011	0.184	0.005	NaN
MLE	50	12	0.5	noise	0.895	0.009	0.131	0.006	NaN
MLE	50	12	0.5	strong	0.897	0.011	0.130	0.007	NaN
Ridge	5	4	0.0	equal	0.782	0.058	0.193	0.030	7.233
Ridge	5	4	0.0	halfnoise	0.671	0.114	0.225	0.024	31.839
Ridge	5	4	0.0	noise	0.611	0.101	0.239	0.018	52.787
Ridge	5	4	0.0	strong	0.733	0.103	0.209	0.035	8.426
Ridge	5	4	0.5	equal	0.821	0.071	0.171	0.037	0.218
Ridge	5	4	0.5	halfnoise	0.684	0.103	0.213	0.030	20.290
Ridge	5	4	0.5	noise	0.762	0.088	0.198	0.037	0.410
Ridge	5	4	0.5	strong	0.833	0.070	0.166	0.034	0.167
Ridge	5	8	0.0	equal	0.851	0.035	0.161	0.021	0.065
Ridge	5	8	0.0	halfnoise	0.686	0.104	0.226	0.023	32.188
Ridge	5	8	0.0	noise	0.752	0.043	0.211	0.021	0.382
Ridge	5	8	0.0	strong	0.823	0.035	0.176	0.021	0.091
Ridge	5	8	0.5	equal	0.931	0.031	0.107	0.023	0.080
Ridge	5	8	0.5	halfnoise	0.751	0.067	0.203	0.021	0.589
Ridge	5	8	0.5	noise	0.821	0.045	0.170	0.024	0.204
Ridge	5	8	0.5	strong	0.872	0.031	0.146	0.018	0.103
Ridge	5	12	0.0	equal	0.862	0.029	0.156	0.016	0.056
Ridge	5	12	0.0	halfnoise	0.706	0.111	0.218	0.021	16.665
Ridge	5	12	0.0	noise	0.806	0.045	0.183	0.024	0.125
Ridge	5	12	0.0	strong	0.812	0.049	0.183	0.026	0.210
Ridge	5	12	0.5	equal	0.972	0.021	0.067	0.023	0.046
Ridge	5	12	0.5	halfnoise	0.819	0.036	0.175	0.018	0.196
Ridge	5	12	0.5	noise	0.903	0.023	0.130	0.016	0.132
Ridge	5	12	0.5	strong	0.913	0.034	0.120	0.023	0.098
Ridge	10	4	0.0	equal	0.777	0.044	0.194	0.017	0.082
Ridge	10	4	0.0	halfnoise	0.658	0.038	0.237	0.009	9.758
Ridge	10	4	0.0	noise	0.656	0.099	0.227	0.020	23.739
Ridge	10	4	0.0	strong	0.775	0.035	0.189	0.016	0.070
Ridge	10	4	0.5	equal	0.834	0.058	0.161	0.034	0.096
Ridge	10	4	0.5	halfnoise	0.691	0.034	0.223	0.009	0.229
Ridge	10	4	0.5	noise	0.714	0.045	0.210	0.016	0.258
Ridge	10	4	0.5	strong	0.824	0.039	0.169	0.018	0.054
Ridge	10	8	0.0	equal	0.836	0.028	0.167	0.014	0.033
Ridge	10	8	0.0	halfnoise	0.684	0.047	0.228	0.016	0.541
Ridge	10	8	0.0	noise	0.733	0.025	0.211	0.008	0.108
Ridge	10	8	0.0	strong	0.767	0.023	0.198	0.009	0.064
Ridge	10	8	0.5	equal	0.915	0.018	0.117	0.012	0.044
Ridge	10	8	0.5	halfnoise	0.760	0.047	0.198	0.017	0.188

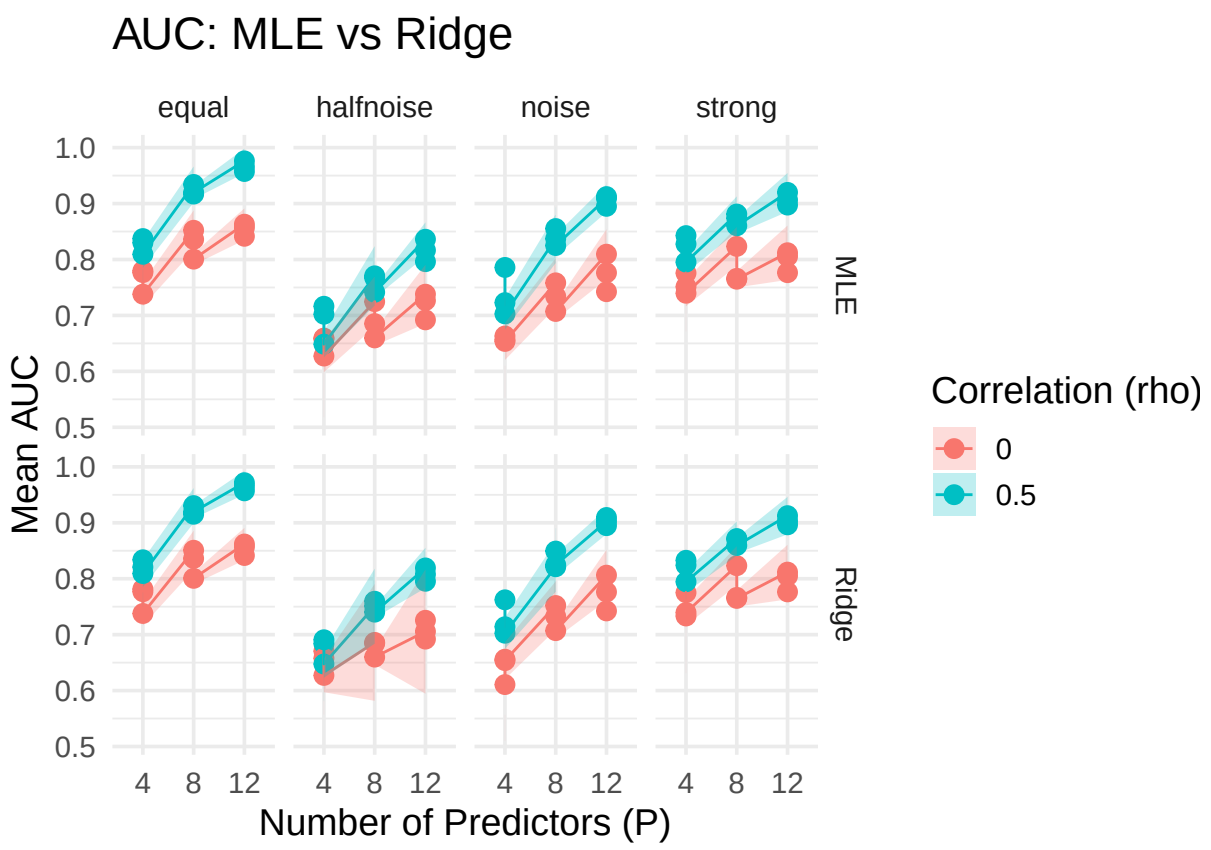
model	EPV	P	rho	beta_pattern	mean_AUC	sd_AUC	mean_Brier	sd_Brier	mean_lambda
Ridge	10	8	0.5	noise	0.849	0.041	0.157	0.022	0.084
Ridge	10	8	0.5	strong	0.871	0.031	0.146	0.017	0.052
Ridge	10	12	0.0	equal	0.857	0.027	0.154	0.015	0.023
Ridge	10	12	0.0	halfnoise	0.726	0.017	0.215	0.007	0.108
Ridge	10	12	0.0	noise	0.776	0.017	0.195	0.007	0.055
Ridge	10	12	0.0	strong	0.805	0.032	0.181	0.014	0.044
Ridge	10	12	0.5	equal	0.965	0.009	0.076	0.009	0.034
Ridge	10	12	0.5	halfnoise	0.808	0.038	0.178	0.018	0.128
Ridge	10	12	0.5	noise	0.910	0.027	0.122	0.019	0.068
Ridge	10	12	0.5	strong	0.903	0.018	0.126	0.013	0.052
Ridge	50	4	0.0	equal	0.738	0.021	0.207	0.008	0.017
Ridge	50	4	0.0	halfnoise	0.627	0.030	0.237	0.005	2.427
Ridge	50	4	0.0	noise	0.654	0.033	0.232	0.008	0.060
Ridge	50	4	0.0	strong	0.740	0.023	0.206	0.008	0.021
Ridge	50	4	0.5	equal	0.809	0.024	0.178	0.011	0.030
Ridge	50	4	0.5	halfnoise	0.648	0.025	0.232	0.006	0.086
Ridge	50	4	0.5	noise	0.703	0.028	0.218	0.008	0.065
Ridge	50	4	0.5	strong	0.795	0.019	0.184	0.008	0.026
Ridge	50	8	0.0	equal	0.801	0.010	0.182	0.004	0.011
Ridge	50	8	0.0	halfnoise	0.660	0.014	0.230	0.004	0.040
Ridge	50	8	0.0	noise	0.707	0.014	0.217	0.004	0.021
Ridge	50	8	0.0	strong	0.765	0.015	0.197	0.006	0.020
Ridge	50	8	0.5	equal	0.920	0.012	0.114	0.009	0.027
Ridge	50	8	0.5	halfnoise	0.740	0.010	0.206	0.003	0.050
Ridge	50	8	0.5	noise	0.825	0.019	0.171	0.009	0.041
Ridge	50	8	0.5	strong	0.859	0.014	0.153	0.007	0.027
Ridge	50	12	0.0	equal	0.841	0.012	0.163	0.006	0.011
Ridge	50	12	0.0	halfnoise	0.692	0.016	0.222	0.005	0.022
Ridge	50	12	0.0	noise	0.742	0.010	0.206	0.004	0.012
Ridge	50	12	0.0	strong	0.776	0.011	0.193	0.004	0.020
Ridge	50	12	0.5	equal	0.957	0.007	0.083	0.007	0.028
Ridge	50	12	0.5	halfnoise	0.795	0.011	0.184	0.005	0.032
Ridge	50	12	0.5	noise	0.895	0.009	0.131	0.006	0.027
Ridge	50	12	0.5	strong	0.896	0.011	0.131	0.007	0.029

```
ggplot(summary_df, aes(x = factor(P), y = mean_AUC,
                       color = model, group = model)) +
  geom_point(size = 3) +
  geom_line() +
  facet_grid(EPV ~ beta_pattern) +
  labs(
    title = "AUC Comparison: MLE vs Ridge Logistic Regression",
    x = "Number of Predictors (P)",
    y = "Mean AUC"
  )
```

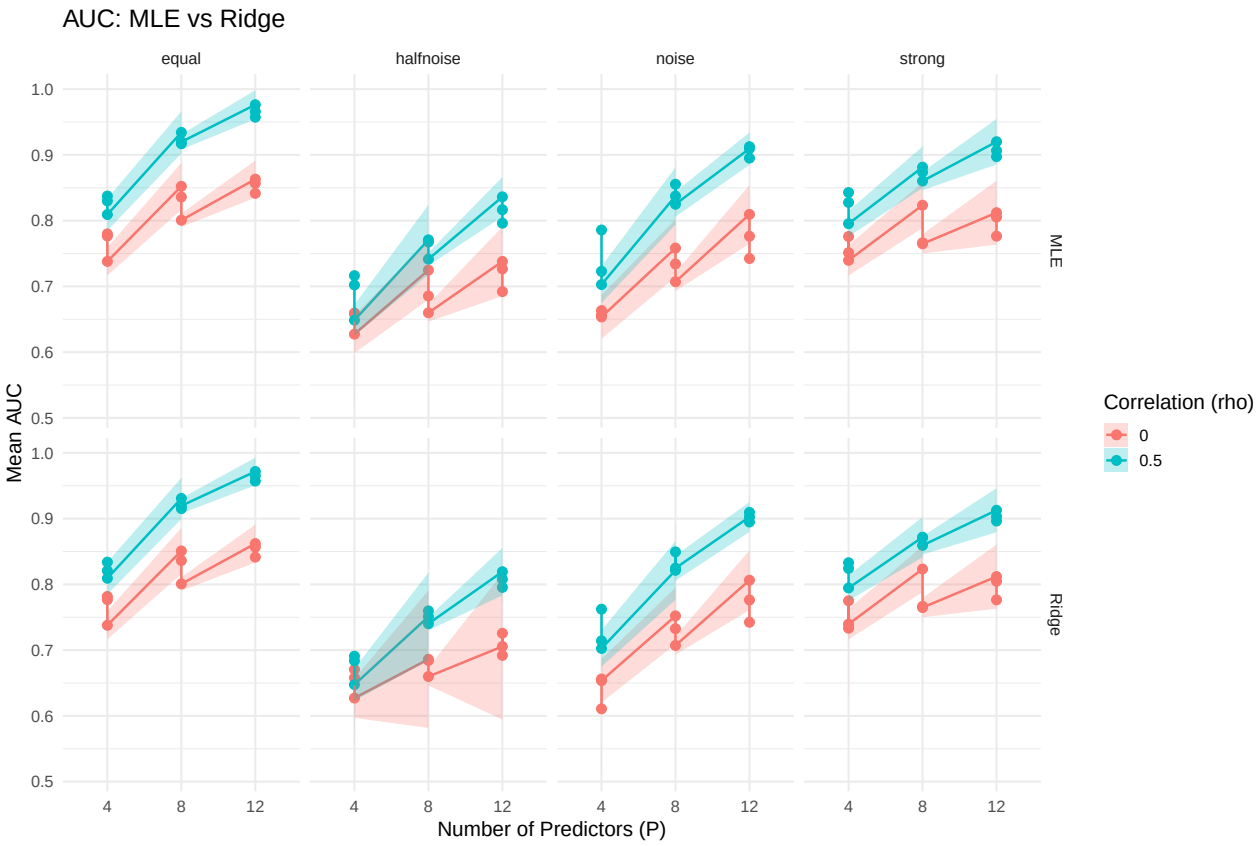



```r
ggplot(summary_df,
       aes(x = factor(P),
           y = mean_AUC,
           color = factor(rho),
           group = factor(rho))) +

  # Points + lines
  geom_point(size = 3) +
  geom_line() +

  # Ribbons for uncertainty
  geom_ribbon(
    aes(ymin = mean_AUC - sd_AUC,
        ymax = mean_AUC + sd_AUC,
        fill = factor(rho)),
    alpha = 0.25,
    color = NA
  ) +

  facet_grid(model ~ beta_pattern) +

  labs(
    title = "AUC: MLE vs Ridge",
    x = "Number of Predictors (P)",
    y = "Mean AUC",
    color = "Correlation (rho)",
    fill  = "Correlation (rho)"
```

```
  ) +
  theme_minimal(base_size = 14)
```




```
ggplot(summary_df,
       aes(x = factor(P),
           y = mean_AUC,
           color = factor(rho),
           group = factor(rho))) +
  geom_point(size = 2.8) +
  geom_line(linewidth = 0.8) +
  geom_ribbon(aes(ymin = mean_AUC - sd_AUC,
                  ymax = mean_AUC + sd_AUC,
                  fill = factor(rho)),
              alpha = 0.25,
              color = NA) +
  facet_grid(model ~ beta_pattern) +
  labs(
    title = "AUC: MLE vs Ridge",
    x = "Number of Predictors (P)",
    y = "Mean AUC",
    color = "Correlation (rho)",
    fill  = "Correlation (rho)"
  ) +
  theme_minimal(base_size = 14)
```


AUC: MLE vs Ridge
equal
halfnoise
noise
strong
MLE
Ridge
Mean AUC
1.0
0.9
0.8
0.7
0.6
0.5
4
8
12
Number of Predictors (P)
Correlation (rho)
0
0.5