

EDA

```
if (!require(pacman)) install.packages("pacman")
```

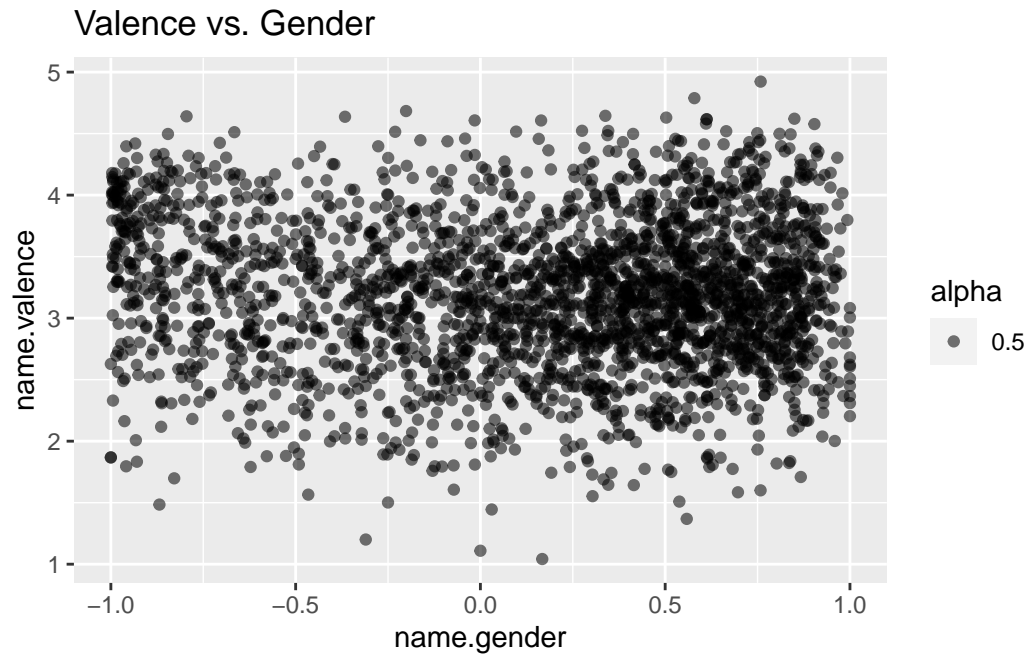
Warning: package 'pacman' was built under R version 4.2.3

```
#> Loading required package: pacman  
pacman::p_load(ChineseNames, tidyverse)
```

```
data(givenname)
```

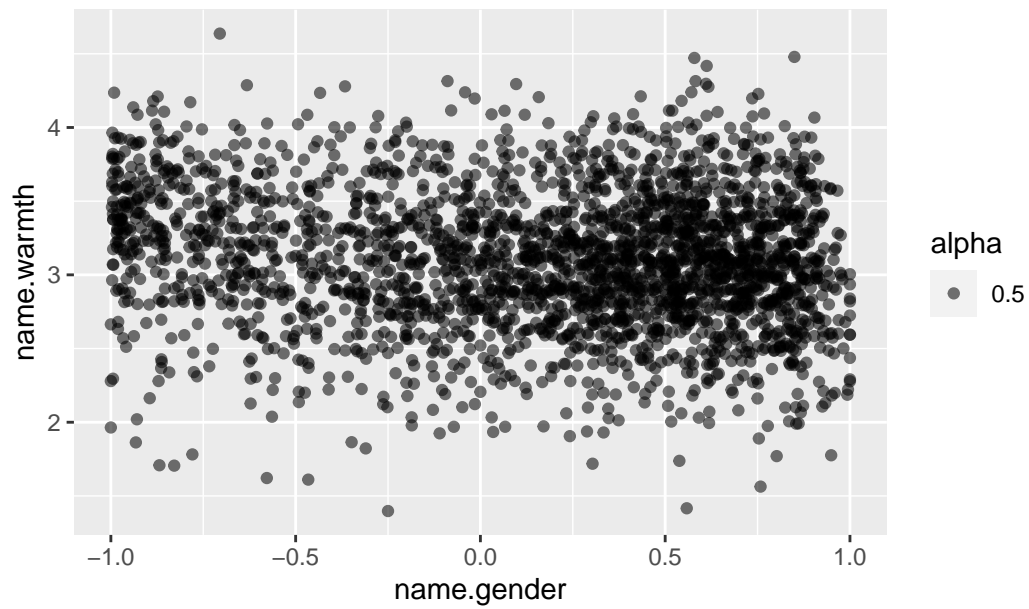
```
set.seed(123)
```

```
ggplot(givenname, aes(x = name.gender, y = name.valence, alpha = 0.5)) +  
  geom_jitter() +  
  labs(  
    title = "Valence vs. Gender"  
  )
```



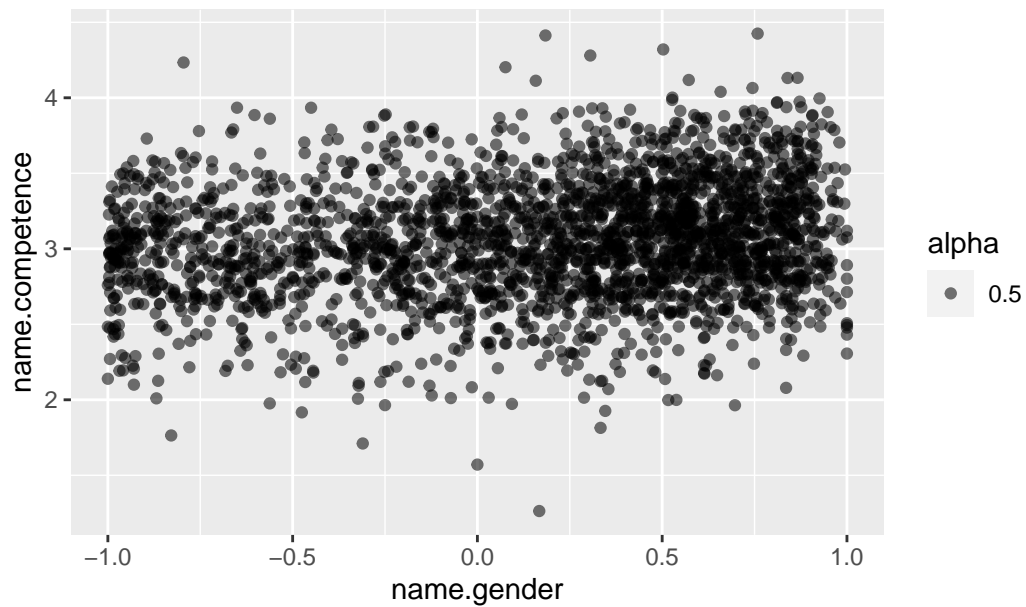
```
ggplot(givenname, aes(x = name.gender, y = name.warmth, alpha = 0.5)) +  
  geom_jitter() +  
  labs(  
    title = "Warmth vs. Gender"  
  )
```

Warmth vs. Gender

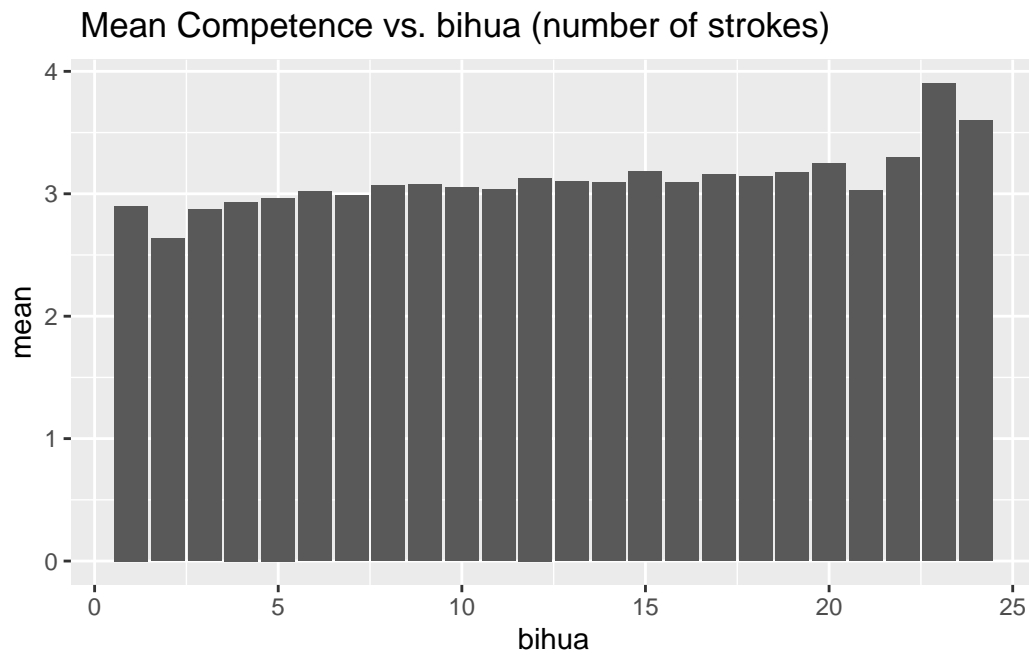


```
ggplot(givenname, aes(x = name.gender, y = name.competence, alpha = 0.5)) +  
  geom_jitter() +  
  labs(  
    title = "Competence vs. Gender"  
  )
```

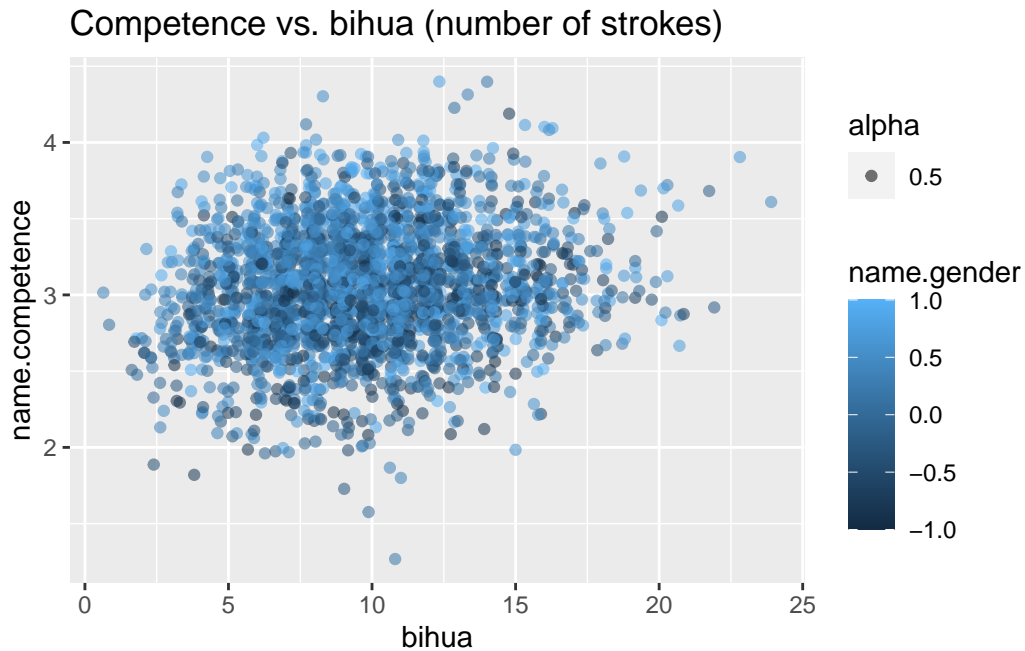
Competence vs. Gender



```
givenname |>
  group_by(bihua) |>
  summarise(mean = mean(name.competence)) |>
  ggplot(aes(x = bihua, y = mean)) +
  geom_col() +
  labs(
    title = " Mean Competence vs. bihua (number of strokes)"
  )
```



```
ggplot(givenname, aes(x = bihua, y = name.competence, color = name.gender, alpha = 0.5))  
geom_jitter() +  
  labs(  
    title = "Competence vs. bihua (number of strokes)"  
  )
```



```
reg1 <- lm(data = givenname, name.competence ~ bihua)
summary(reg1)
```

Call:

```
lm(formula = name.competence ~ bihua, data = givenname)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.77561	-0.24221	-0.00902	0.25779	1.30769

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.891893	0.022228	130.099	< 2e-16 ***
bihua	0.016702	0.002157	7.745	1.36e-14 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3815 on 2612 degrees of freedom

Multiple R-squared: 0.02245, Adjusted R-squared: 0.02207

F-statistic: 59.98 on 1 and 2612 DF, p-value: 1.361e-14