

```
!pip install plotnine pandas -q
```

```
from plotnine import ggplot, aes, geom_point, geom_smooth, facet_grid, labs, theme_minimal, theme, element_text, facet_wrap
from plotnine.data import diamonds
import pandas as pd
from datetime import datetime
print("Arquivo compilado em:", datetime.now().strftime("%d/%m/%Y %H:%M:%S"))

diamonds.head()
```

Arquivo compilado em: 30/10/2025 15:43:05

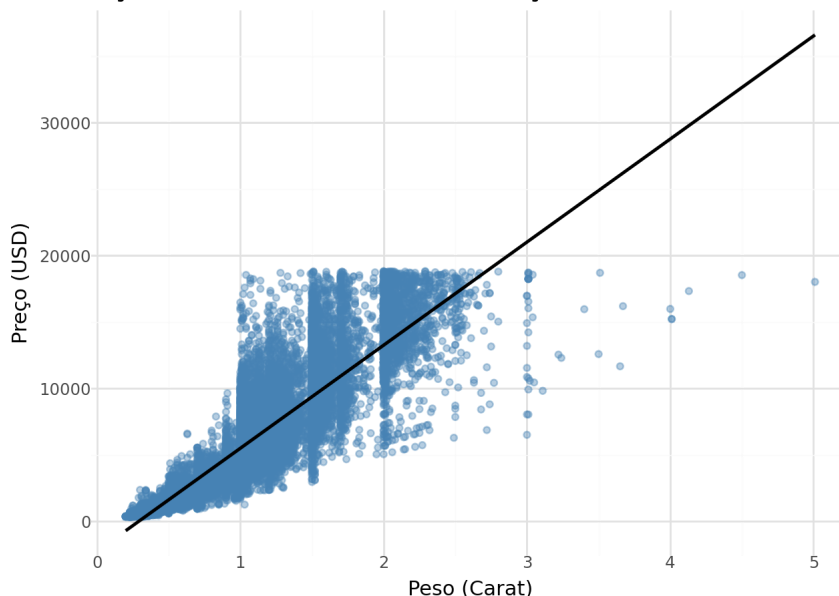
	carat	cut	color	clarity	depth	table	price	x	y	z	
0	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43	
1	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31	
2	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31	
3	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63	
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75	

Próximas etapas:

[Gerar código com diamonds](#)
[New interactive sheet](#)

```
plot01 = (
    ggplot(diamonds, aes(x='carat', y='price'))
    + geom_point(alpha=0.4, color='steelblue')
    + geom_smooth(method='lm', color='black', se=False)
    + labs(
        title='Relação entre Peso (Carat) e Preço (Price) dos Diamantes',
        x='Peso (Carat)',
        y='Preço (USD)'
    )
    + theme_minimal()
    + theme(
        plot_title=element_text(weight='bold', size=14)
    )
)
plot01
```

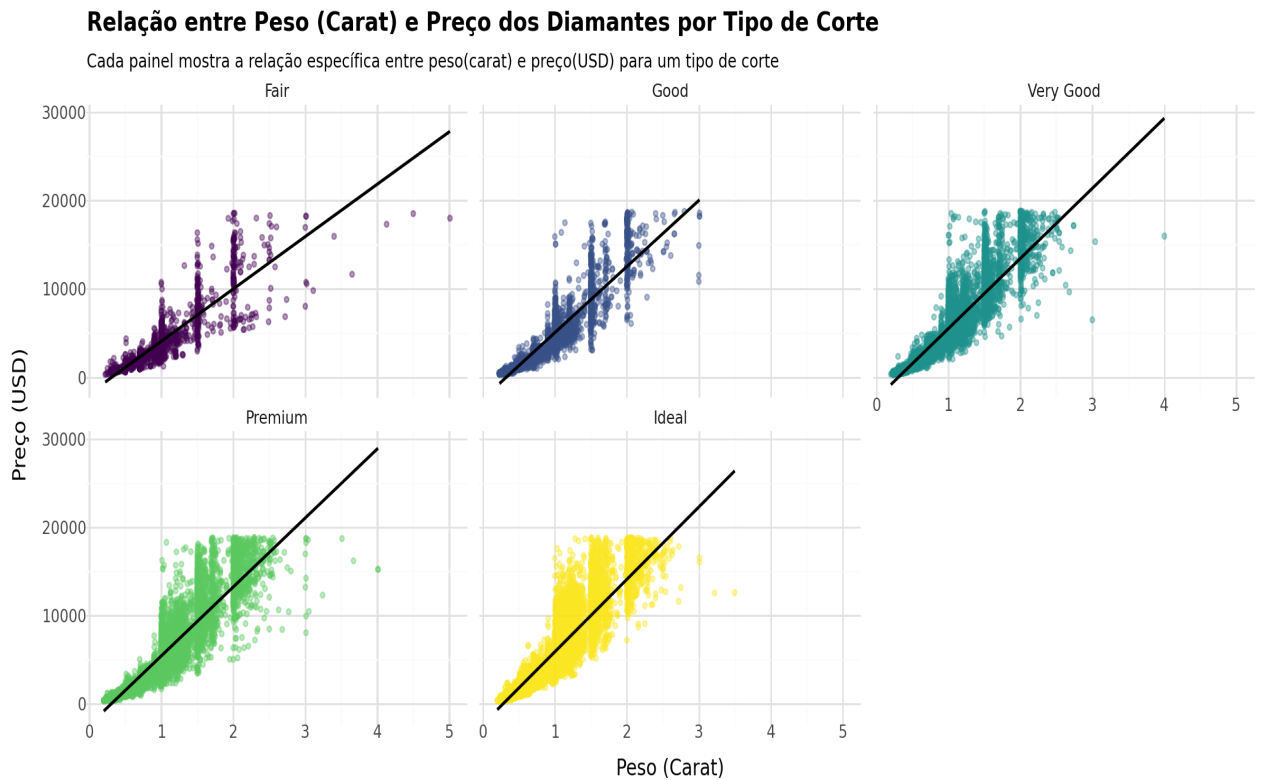
**Relação entre Peso (Carat) e Preço (Price) dos Diamant**



A análise do gráfico mostra uma clara relação positiva entre o peso (carat) e o preço dos diamantes: quanto maior o peso, maior tende a ser o preço. No entanto, essa relação não é perfeitamente linear — observa-se uma dispersão crescente conforme o peso aumenta, indicando que outros fatores (como o corte, representado pelas cores) também influenciam o preço final. Diamantes com cortes de melhor qualidade, como Ideal e Premium, geralmente apresentam preços mais altos em comparação aos de corte Fair para o mesmo peso.

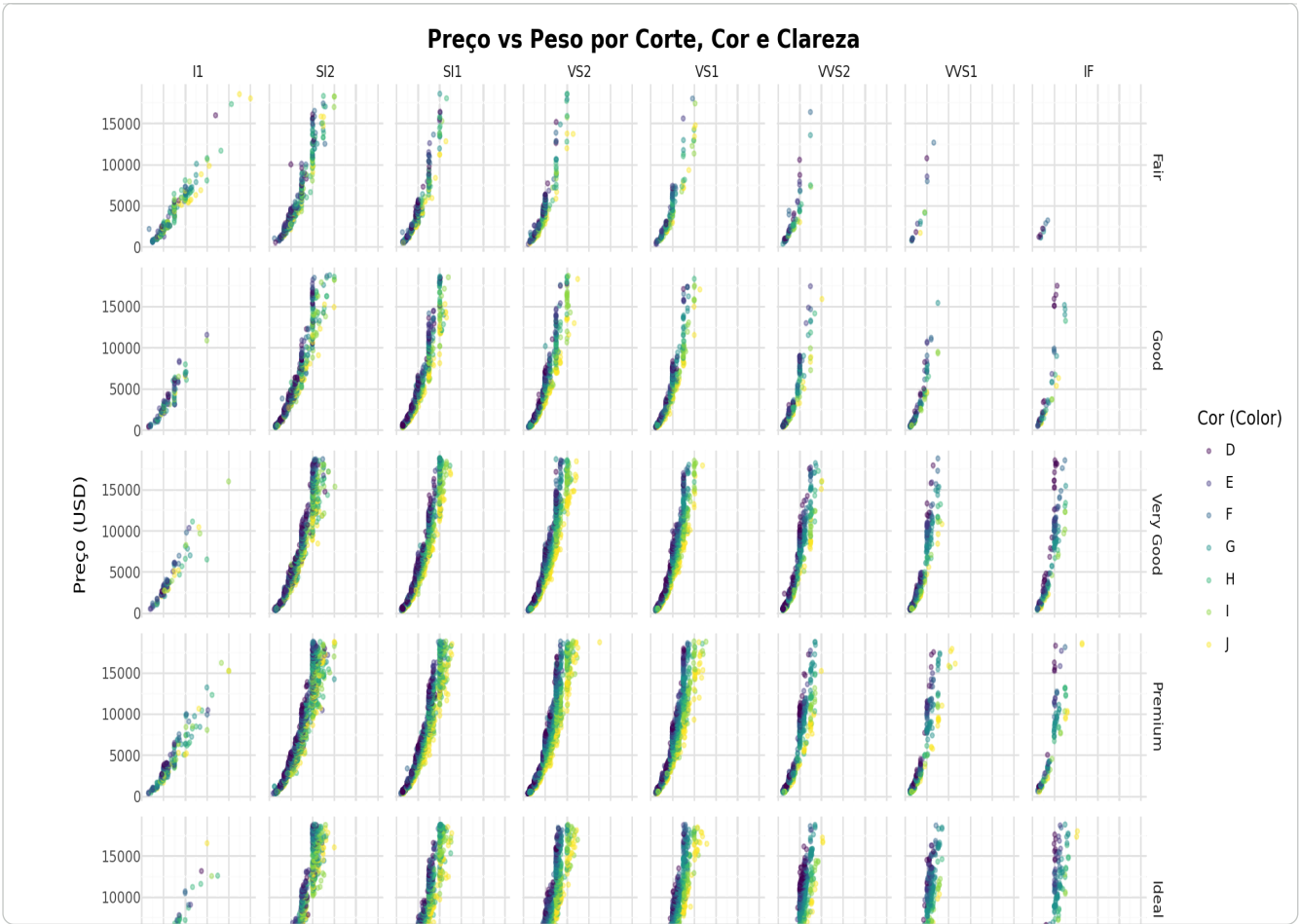
```
plot02 = (
    ggplot(diamonds, aes(x='carat', y='price', color='cut')) +
```

```
geom_point(alpha=0.4, size=1.2, show_legend=False) + # pontos sem legenda repetida
geom_smooth(method='lm', color='black', se=False) + # linha de tendência
facet_wrap('~cut') + # cria um gráfico por tipo de corte
labs(
  title='Relação entre Peso (Carat) e Preço dos Diamantes por Tipo de Corte',
  subtitle='Cada painel mostra a relação específica entre peso(carat) e preço(USD) para um tipo de corte',
  x='Peso (Carat)',
  y='Preço (USD)'
) +
theme_minimal(base_size=12) +
theme(
  plot_title=element_text(weight='bold', size=14),
  plot_subtitle=element_text(size=10),
  figure_size=(12,6)
)
)
plot02
```



Observando os gráficos, nota-se que a relação positiva entre o peso (carat) e o preço (price) se mantém em todos os tipos de corte. No entanto, a intensidade dessa relação (inclinação da linha) e a dispersão dos preços variam. Cortes de melhor qualidade, como Ideal e Premium, tendem a apresentar preços mais altos e mais consistentes para o mesmo peso, enquanto cortes inferiores, como Fair e Good, mostram maior variação e preços médios mais baixos. Isso indica que, embora o peso seja um fator importante no preço do diamante, o tipo de corte exerce um papel relevante no valor final.

```
plot03 = ggplot(diamonds, aes(x='carat', y='price', color='color')) + \
  geom_point(alpha=0.4, size=1) + \
  facet_grid('cut ~ clarity') + \
  labs(
    title='Preço vs Peso por Corte, Cor e Clareza',
    x='Peso (Carat)',
    y='Preço (USD)',
    color='Cor (Color)'
  ) + \
  theme_minimal() + \
  theme(
    plot_title=element_text(weight='bold', size=14),
    figure_size=(12, 8),
    axis_text_x=element_text(rotation=45, hjust=1)
  )
)
plot03
```



A principal dificuldade foi visualizar simultaneamente três variáveis categóricas (corte, cor e clareza) junto às variáveis numéricas (carat e price). Como o conjunto de dados é grande, o gráfico pode ficar muito denso e de difícil leitura se mostrarmos tudo de uma vez. Foi necessário reduzir a amostra e usar facetas e cores para tornar as relações compreensíveis. Mesmo assim, há limitações — o excesso de categorias reduz a clareza visual. Apesar das dificuldade, é possível perceber que a relação entre peso (carat) e preço (USD) se mantém positiva em todas as combinações, mas a força dessa relação e o nível de preço variam conforme o corte, a cor e a clareza. Diamantes com melhor corte (Ideal, Premium), cor mais clara (D, E) e maior clareza (IF, VVS1) apresentam valores mais altos para o mesmo peso, indicando que essas características agregam valor adicional ao preço. Já os diamantes com cortes e cores inferiores tendem a apresentar maior dispersão e preços menores, mesmo com o mesmo peso.

```
from datetime import datetime
print("Relatório compilado em:", datetime.now().strftime("%d/%m/%Y %H:%M:%S"))
```

Relatório compilado em: 30/10/2025 15:43:27