

07

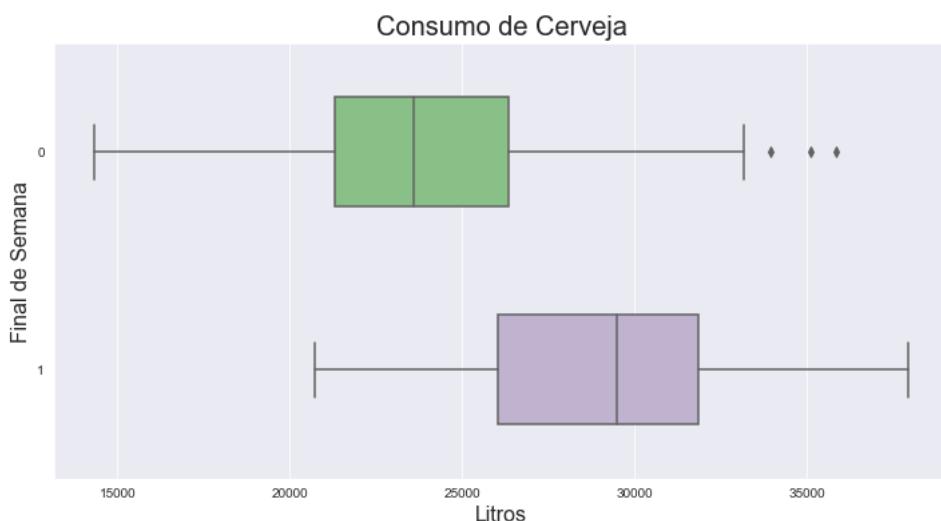
## Comparando distribuições

Em nossa última aula, aprendemos a construir *box plots* com o cruzamento das informações de duas variáveis de nosso *dataset*. Esta técnica de visualização nos ajuda a identificar diferenças na distribuição de uma variável, quando separada em determinadas categorias.

Utilizando o nosso *dataset*, vamos executar o código abaixo para obter o *box plot* do consumo de cerveja, segundo a variável final de semana:

```
ax = sns.boxplot(y='fds', x='consumo', data=dados, orient='h', width=0.5)
ax.figure.set_size_inches(12, 6)
ax.set_title('Consumo de Cerveja', fontsize=20)
ax.set_xlabel('Litros', fontsize=16)
ax.set_ylabel('Final de Semana', fontsize=16)
ax
```

Note que é o mesmo *box plot* construído em nossa aula, mudando somente a orientação (`orient='h'`):



Avalie as alternativas abaixo e selecione as corretas. Lembre-se que a variável "Final de Semana" assume o valor 0 (zero) para dias de semana e 1 (um) para finais de semana.

Selecionar 2 alternativas

**A** Quando analisamos o consumo de cerveja fora dos finais de semana, notamos alguns valores discrepantes, que devem ser investigados com maior atenção.

**B** É possível inferir, através do gráfico, que o consumo de cerveja nos finais de semana é mais elevado.

**C** É possível afirmar que as duas distribuições são perfeitamente simétricas.

**D** É possível inferir, através do gráfico, que o consumo de cerveja nos finais de semana é menos intenso.

