Why does Kruskal's algorithm work?

We summarize the proof by contradiction:

- We let T^* be a MST different from T
- We took any edge e from T^* that wasn't in T, and we added it to T
- We argued this created a cycle γ and that e had to be the heaviest edge in γ
- We argued there was an edge f in γ that isn't in T^* and we could add it to $T^* - e$ and get a spanning tree that had smaller weight
- Hence we obtained a contradiction