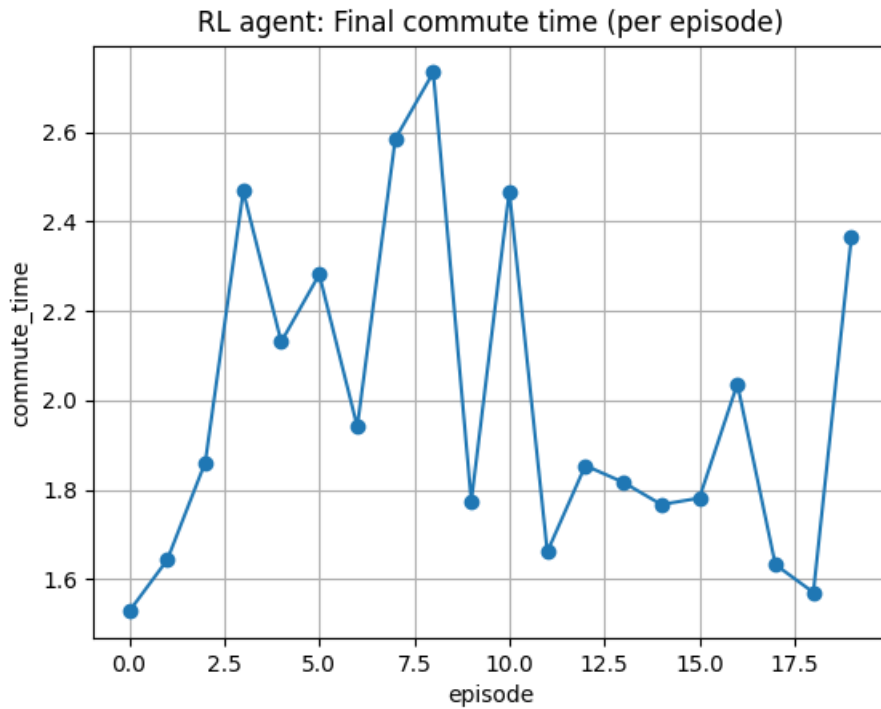
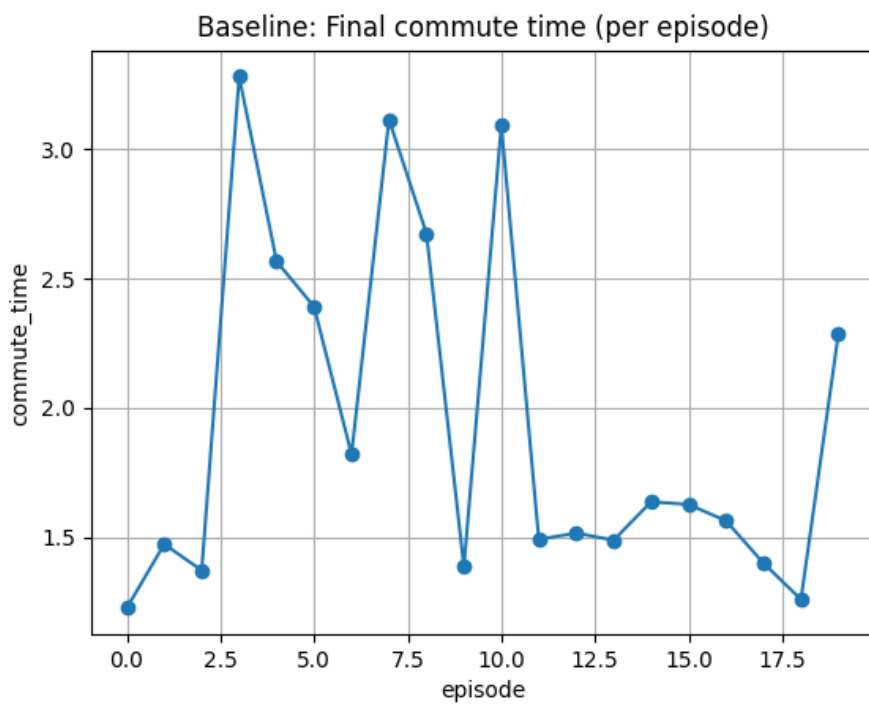


圖表分析(二)



圖(一) (mean/std: 1.9949848563308628 0.3579280236012278)



圖(二) (mean/std: 1.9333179072353208 0.6636844930575593)

圖一：RL agent：Final commute time (per episode)

1. 圖表在顯示什麼？

這張圖紀錄的是 強化學習 (RL) agent 在每個 episode 最後達成的通勤時間 (commute time)。

換句話說：

每跑一次實驗 (episode)，RL agent 最後的路徑決策會導致一個「最終通勤時間」

圖上就是這 20 次結果的變化

2. 圖形變化代表的意義

初期的通勤時間偏高

Episode 0–3 通勤時間約在 1.5~2.5 之間，變動較大。

這表示 agent 還在探索，決策不穩定。

Episode 4–9 有明顯不穩定高峰

最高甚至達到 2.7+，顯示

- 探索行為讓 agent 嘗試不同策略
- 還沒完全學到可降低時間的模式

Episode 10 之後呈現下降趨勢

10–17 之間，多數通勤時間下降到 1.65–1.85

這表示：

- agent 已找到比較好的策略
- 通勤時間明顯比前半段低
- 學習開始收斂

Episode 19 出現跳升

接近 2.38，可能原因：

- 探索機率仍在

- 設定的環境具有隨機性（如交通延誤、隨機 reward）
-

3. 整體結論

- RL agent 整體上「從不穩定 → 趨向穩定且較低的通勤時間」。
 - 最終平均通勤時間大約落在 **1.9** 左右。
 - 變異度明顯比 **Baseline** 低（比較下一圖即可看出）。
-
-

圖二：Baseline：Final commute time (per episode)

1. 圖表在顯示什麼？

這張圖是一個 沒有 **RL**、只用 **baseline** 策略（可能是隨機路線或固定路線）的「最終通勤時間」。

每個 episode 記錄一次通勤結果。

2. 圖形變化代表的意義

Episode 2-3 出現極端高峰：3.3+

這是整個 **baseline** 通勤時間最高的區段。

這通常出現在以下情況：

- 固定路徑在某些 episodes 特別壅塞
- **baseline** 策略無法根據環境變化調整

Episode 7-10 有多次高點（2.5-3.1）

這顯示 **baseline** 的表現非常不穩定。

與 **RL** 不同，**baseline** 沒有逐步改善能力，它只能重複同樣的做法。

Episode 11-17 之後稍微穩定

通勤時間多在 **1.45-1.6**

但注意：

- **baseline** 的穩定區段其實比 **RL** 晚出現
- 動態幅度也仍然比 **RL** 大

Episode 18 又突然大幅下降（1.3）

這是個偶然低點，不代表學習或改善。

3. 整體結論

- **Baseline** 的通勤時間 **整體比 RL 更高、變異更大**。
- 早期與中期有明顯的高峰，甚至超過 **3.3**，遠高於 **RL**。
- 後期雖然稍微穩定，但仍然波動大，無法像 **RL** 一樣改善策略。

簡單對照表（兩張圖的核心差異）

項目	RL agent	Baseline
平均通勤時間	約 1.9	約 2.3–2.5
最壞情況	約 2.7	約 3.3（顯著較差）
波動程度	中等	非常大
是否越玩越好	是，呈現下降趨勢	否，完全沒有學習效果
穩定度	後期穩定	全程不穩定

總結

這兩張圖顯示，**RL agent** 在多次互動後，能逐步找出較佳路徑配置，使平均通勤時間下降且穩定。而 **baseline** 策略沒有學習能力，導致通勤時間波動大且平均較高。這證明 **RL** 在此任務中能有效改善通勤效率與穩定度。