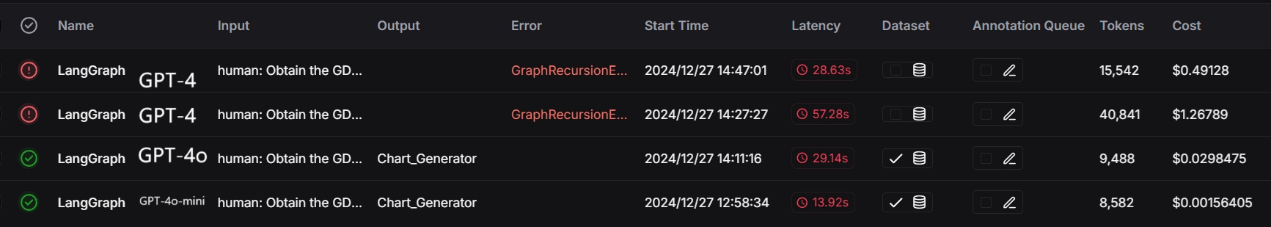
# LangSmith监测结果



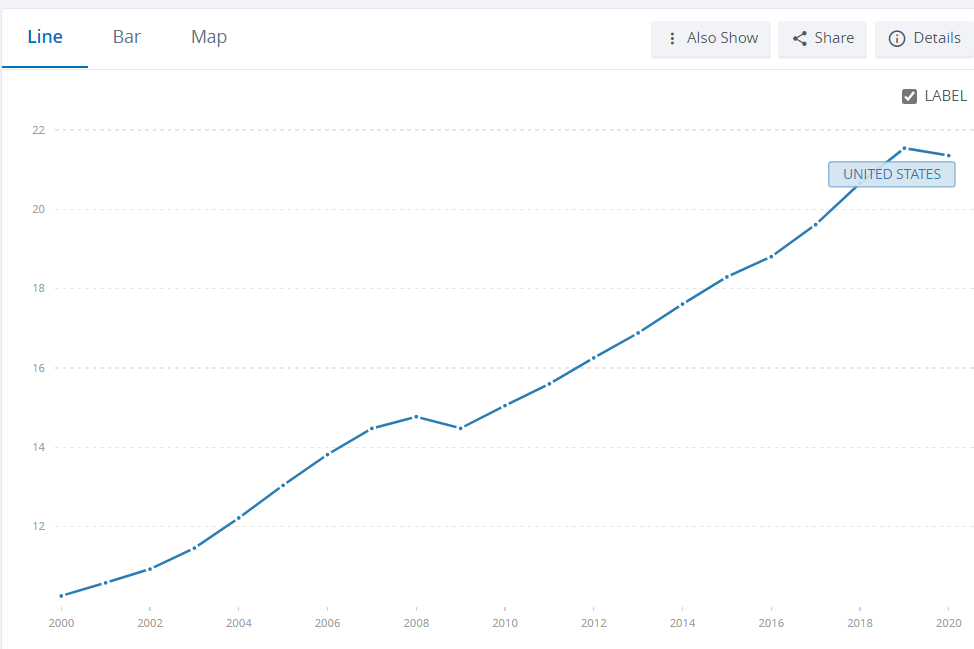
GPT-4o的结果最好，和世界银行结果相符合。

GPT-4o-mini的结果一般，检索的数据不是很好，GDP是以2012 美元计。

GPT-4运行时虽然输出了统计图，但是生成质量较差，达到设置最大递归限制。检查输出日志发现，没有按照提示词生成终止符FINAL ANSWER。需要调整提示词，要求在完成任务后仅输出FINAL ANSWER。输出结果最差，有部分年份数据缺失。

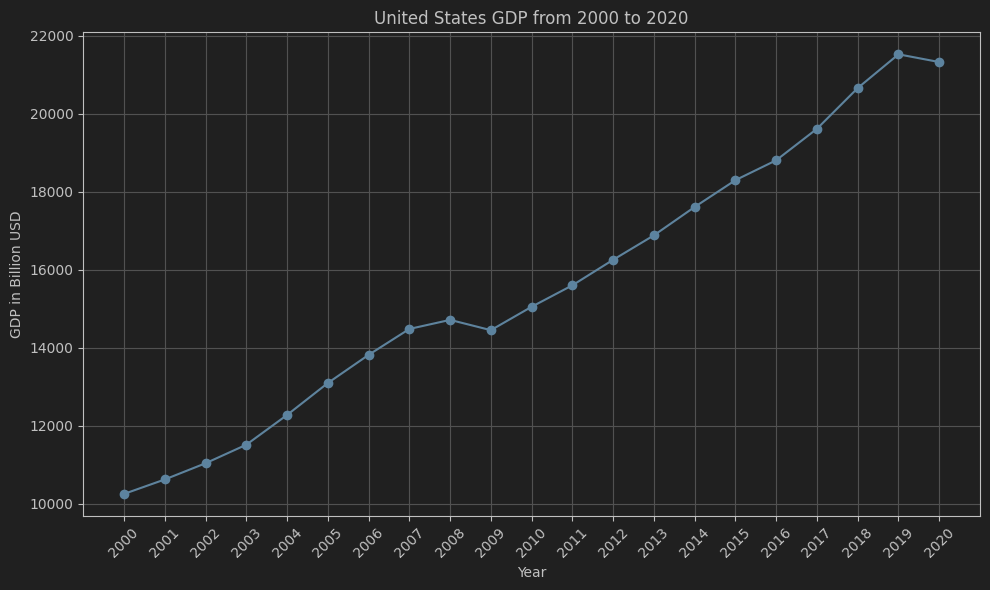
# 世界银行显示结果

[GDP (current US$) - United States | Data](https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?contextual=default&end=2020&locations=US&start=2000)

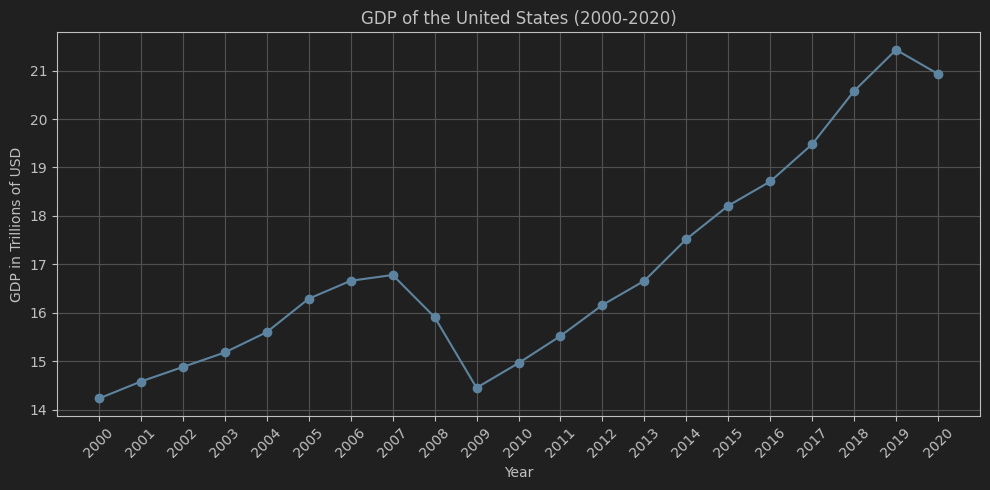


# GPT-4o结果

和世界银行的数据相吻合



# GPT-4o结果



# GPT4结果

