```
import excel "F:\Users\admin\Desktop\CUHK (DSPS)\CUHK tourist data\Courseware grant\hk_visitors_sampledata.xlsx", sheet("data") firstrow clear

tsset time

tsline total_va, name(graph1, replace) title("{bf: Total visitor arrivals in HK, Jan 2016-Dec 2020}") ytitle("") ttitle("")

* Descriptive statistics

gen period = "1. Normal_times" if inrange(time, td(1jan2016), td(31may2019))

replace period = "2. Ebill_times" if inrange(time, td(1jun2019), td(31jan2020))

replace period = "3. COVID-19 times" if inrange(time, td(1feb2020), td(31dec2020))

sum total_va if period=="1. Normal_times", d

sum total_va if period=="2. Ebill_times", d

sum total_va if period=="3. COVID-19 times", d

graph box total_va, over(period) noout ytitle("") name(graph2, replace)

* Confidence interval

gen Mainland_proportion = total_va_mainland_china/total_va

ci means Mainland_proportion if period=="1. Normal_times"

ci means Mainland_proportion if period=="2. Ebill_times"

ci means Mainland_proportion if period=="3. COVID-19 times"

twoway (scatter Mainland_proportion time) (lfitci Mainland_proportion time if period=="1. Normal_times", ciplot(rline)), ytitle("") name(graph3, replace)

twoway (scatter Mainland_proportion time) (lfitci Mainland_proportion time if period=="2. Ebill_times", ciplot(rline)), ytitle("") name(graph4, replace)

twoway (scatter Mainland_proportion time) (lfitci Mainland_proportion time if period=="3. COVID-19 times", ciplot(rline)), ytitle("") name(graph5, replace)
```