A. Matlab code programmed

A.1 Main body: algorithm "*Main.m*"

Developed to perform all the tasks specified in the report. Helped by the usage of some functions specified following.

```
% Read data from excel
filename = input ('please input your filename:','s');
[num,txt,raw] = xlsread(filename);

servicerecord = raw;

% Data Cleaning
forrecord_index = 1:size(servicerecord,1)
gocard_index(record_index,:) =
iscellstr(servicerecord(record_index,7));
end
loc_nongocard = find (gocard_index == 0);
servicerecord (loc_nongocard,:) = [];

loc_type = find (cell2mat(servicerecord(:,11))==1);
SRtemp = servicerecord(loc_type,:);

loc_nan = cellfun(@(x) any(isnan(x)),SRtemp(:,9));
SRtemp(loc_nan,:) = [];

Cleaned_Data = SRtemp;
CleanData_Test = SRtemp;
%xlswrite('datacleaning',SRtemp)

%Time and Date to num (col 8)
ind = size (Cleaned_Data,1);
fori = 1:ind
D=Cleaned_Data{i,2};
H=Cleaned_Data{i,8};
d=datenum(D);
h=datenum(H);
inici=h-d;
CleanData_Test{i,8} = inici;
D2 = Cleaned_Data{i,2};
dd2 = datenum(D2);
```

```
end
%Time and Date to num (col 9)
ind2 = size (Cleaned_Data,1);
fori = 1:ind2
D=Cleaned_Data{i,2};
H=Cleaned_Data{i,9};
d=datenum(D);
h=datenum(H);
final=h-d;
CleanData_Test{i,9} = final;
end
fori = 1:ind
D=Cleaned_Data{i,2};
d=datenum(D,'dd/mm/yyyy');
CleanData_Test{i,2} = d;
end
Cleaned_Data = CleanData_Test;

CCSize = size (Cleaned_Data,1);
Route = Classify(Cleaned_Data(:,3),CCSize,Cleaned_Data);

Data109 = Route{1,1};
%Data412 = Route{2,1};

%StationBoa = cell(10^5,7);
%StationAli = cell(10^5,7);

rwsD109=size(Data109,1);
CS1=Classify(Data109(:,2),rwsD109,Data109);

fori = 1:(size(CS1,1))
rwsCS1 = size (CS1{i,1},1);
CS2{i,1} = Classify(CS1{i,1}(:,4),rwsCS1,CS1{i,1});
end



fori = 1:(size(CS2,1))
for j = 1:(size(CS2{i,1},1))
rwsCS2 = size (CS2{i,1}{j,1},1);
CS3{i,1}{j,1} = Classify(CS2{i,1}{j,1}(:,5),rwsCS2,CS2{i,1}{j,1});
end
end

CS3Boas = Boa(CS3);
CS3Alis = Ali(CS3);

fori = 1:(size(CS3Boas,1))
for j = 1:(size(CS3Boas{i,1},1))
for k = 1:(size(CS3Boas{i,1}{j,1},1))
```

```
CS4{i,1}{j,1}{k,1} =
[CS3Boas{i,1}{j,1}{k,1};CS3Alis{i,1}{j,1}{k,1}];
end
end
end

%End of Classifying and clustering stage
%Sort rows in ascending order to differentiate different bus groups

fori = 1:(size(CS4,1))
for j = 1:(size(CS4{i,1},1))
for k = 1:(size(CS4{i,1}{j,1},1))
aux = sortrows(CS4{i,1}{j,1}{k,1},4);
CS4v2{i,1}{j,1}{k,1} = aux;
end
end
end

%ITERATIVE METHOD to find the optimal number of clusters

fori = 1:(size(CS4,1))
for j = 1:(size(CS4{i,1},1))
for k = 1:(size(CS4{i,1}{j,1},1))
nC=0;
a=pdist(cell2mat(CS4v2{i,1}{j,1}{k,1}(:,4)))
n=size(CS4v2{i,1}{j,1}{k,1}(:,4),1);
nT=n;
m=1;
ind=1;
cz=0;
cT=0;

whileind<=(size(a,2))
if a(ind)<= 0.014
cz=cz+1;
else
end
if m<(n-1)
ind=ind+1;
m=m+1;
else
m=1;
n=n-1;
ind=ind+1;
ifcz>0
cT=cT+1;
cz=0;
else
cz=0;
end
```

```
end
end

nC = nT - cT;
ifnC==1

nC=1.8;
end

A=CS4{i,1}{j,1}{k,1}(:,4);
B=sort(cell2mat(A));
C=clusterdata(B,nC)
CS4v2{i,1}{j,1}{k,1}(:,9) = num2cell(C);

end
end
end


%splitting clusters into bus trips: It creates a new data division
fori = 1:(size(CS4v2,1))
for j = 1:(size(CS4v2{i,1},1))
for k = 1:(size(CS4v2{i,1}{j,1},1))
siz_rows=size(CS4v2{i,1}{j,1}{k,1},1);
CS5{i,1}{j,1}{k,1}=Classify(CS4v2{i,1}{j,1}{k,1}(:,9),siz_rows,CS4v2
{i,1}{j,1}{k,1});
end
end
end

%Joining records of same stations together
fori = 1:(size(CS5,1))
for j = 1:(size(CS5{i,1},1))
for k = 1:(size(CS5{i,1}{j,1},1))
for l = 1:(size(CS5{i,1}{j,1}{k,1},1))
AUX=Classify(CS5{i,1}{j,1}{k,1}{l,1}(:,6),size(CS5{i,1}{j,1}{k,1}{l,
1},1),CS5{i,1}{j,1}{k,1}{l,1});
CS6{i,1}{j,1}{k,1} = list(AUX);
CS6{i,1}{j,1}{k,1} = sortrows(CS6{i,1}{j,1}{k,1},4);
end
          AUX=[];
end
end
end

%Calculations: Travel Time, bus line 109
CS7 = CS6;
fori = 1:(size(CS6,1))
for j = 1:(size(CS6{i,1},1))
```

```
for k = 1:(size(CS6{i,1}{j,1},1))
a = strncmpi(CS6{i,1}{j,1}{k,1}{1,3},'Inbound',3);
if CS6{i,1}{j,1}{k,1}{1,2} == 109
for l = 1:(size(CS6{i,1}{j,1}{k,1},1))
aux = strncmpi(CS6{i,1}{j,1}{k,1}{l,5},'UQ Lakes',8);
aux2 = strncmpi(CS6{i,1}{j,1}{k,1}{l,5},'Cultural Centre',8);
aux3 = strncmpi(CS6{i,1}{j,1}{k,1}{l,5},'Adelaide St',8);
if a == 1
if aux == 1
CS7{i,1}{j,1}{k,1}{1,12} = CS6{i,1}{j,1}{k,1}{l,4};
elseif aux2 == 1
CS7{i,1}{j,1}{k,1}{1,13} = CS6{i,1}{j,1}{k,1}{l,4};
elseif aux3 == 1
CS7{i,1}{j,1}{k,1}{1,14} = CS6{i,1}{j,1}{k,1}{l,4};
end
else
if aux3 == 1
CS7{i,1}{j,1}{k,1}{1,12} = CS6{i,1}{j,1}{k,1}{l,4};
elseif aux2 == 1
CS7{i,1}{j,1}{k,1}{1,13} = CS6{i,1}{j,1}{k,1}{l,4};
elseif aux == 1
CS7{i,1}{j,1}{k,1}{1,14} = CS6{i,1}{j,1}{k,1}{l,4};
end
end
%Dwell time
adw = CS7{i,1}{j,1}{k,1}{l,6};
bdw = CS7{i,1}{j,1}{k,1}{l,4};
CS7{i,1}{j,1}{k,1}{l,15} = adw-bdw;
%travel Time#2
CS7{i,1}{j,1}{k,1}{2,12} = CS7{i,1}{j,1}{k,1}{1,13} -
CS7{i,1}{j,1}{k,1}{1,12};
CS7{i,1}{j,1}{k,1}{2,13} = CS7{i,1}{j,1}{k,1}{1,14} -
CS7{i,1}{j,1}{k,1}{1,13};
CS7{i,1}{j,1}{k,1}{2,14} = CS7{i,1}{j,1}{k,1}{1,14} -
CS7{i,1}{j,1}{k,1}{1,12};
CS7{i,1}{j,1}{k,1}{1,19} =  sum (cell2mat(CS7{i,1}{j,1}{k,1}(:,7)));
end

else%412 bus calculations:
for l = 1:(size(CS6{i,1}{j,1}{k,1},1))
aux = strncmpi(CS6{i,1}{j,1}{k,1}{l,5},'University',8);
aux2 = strncmpi(CS6{i,1}{j,1}{k,1}{l,5},'Adelaide St',8);
if a == 1
if aux == 1
CS7{i,1}{j,1}{k,1}{1,12} = CS6{i,1}{j,1}{k,1}{l,4};
elseif aux2 == 1
CS7{i,1}{j,1}{k,1}{1,13} = CS6{i,1}{j,1}{k,1}{l,4};
end
else
if aux2 == 1
CS7{i,1}{j,1}{k,1}{1,12} = CS6{i,1}{j,1}{k,1}{l,4};
elseif aux == 1
```

```
CS7{i,1}{j,1}{k,1}{1,13} = CS6{i,1}{j,1}{k,1}{l,4};
end
end

%Dwell time
adw = CS7{i,1}{j,1}{k,1}{l,6};
bdw = CS7{i,1}{j,1}{k,1}{l,4};
CS7{i,1}{j,1}{k,1}{l,14} = adw-bdw;
%travel Time#2
CS7{i,1}{j,1}{k,1}{2,12} = CS7{i,1}{j,1}{k,1}{1,13} -
CS7{i,1}{j,1}{k,1}{1,12};
CS7{i,1}{j,1}{k,1}{1,18} =  sum (cell2mat(CS7{i,1}{j,1}{k,1}(:,7)));
end
end

%Dwell time#2
CS7{i,1}{j,1}{k,1}{1,15} =  sum
(cell2mat(CS7{i,1}{j,1}{k,1}(:,14)));
if CS7{i,1}{j,1}{k,1}{2,12} ~=0
CS7{i,1}{j,1}{k,1}{1,16} = (CS7{i,1}{j,1}{k,1}{2,12})-
CS7{i,1}{j,1}{k,1}{1,15};
else
CS7{i,1}{j,1}{k,1}{1,16} = [];
end

%N° of passengers inside the bus
for l = 1
CS7{i,1}{j,1}{k,1}{l,17} = ((CS7{i,1}{j,1}{k,1}{l,7})-
(CS7{i,1}{j,1}{k,1}{l,9}));
end
for l = 2:(size(CS7{i,1}{j,1}{k,1},1))
CS7{i,1}{j,1}{k,1}{l,17} = CS7{i,1}{j,1}{k,1}{l-1,17} +
((CS7{i,1}{j,1}{k,1}{l,7})-(CS7{i,1}{j,1}{k,1}{l,9}));
end
%Count if bus exceeds limit of 62 passengers
for l = 1:(size(CS7{i,1}{j,1}{k,1},1))
aux_lim = 0;
if CS7{i,1}{j,1}{k,1}{l,17}>61
CS7{i,1}{j,1}{k,1}{2,18} = 'limit exceeded';
CS7{i,1}{j,1}{k,1}{3,18} = CS7{i,1}{j,1}{k,1}{l,18};
CS7{i,1}{j,1}{k,1}{4,18} = CS7{i,1}{j,1}{k,1}{l,5};
aux_lim = 1;

elseifaux_lim == 0
CS7{i,1}{j,1}{k,1}{2,18} = 'limit not exceeded';
CS7{i,1}{j,1}{k,1}{3,18} = max(cell2mat(CS7{i,1}{j,1}{k,1}(:,17)));
end
end

%end of triple loop (i,j,k variables):
end
```

```
end
end

list_bus = cell(1,13);
fori = 1:(size(CS7,1))
for j = 1:(size(CS7{i,1},1))
for k = 1:(size(CS7{i,1}{j,1},1))
list_bus = createlist(list_bus,CS7{i,1}{j,1}{k,1});
end
end
end

fori = 1:(size(list_bus,1))
iflist_bus{i,11} < 2
for j = 1:(size(list_bus,2))
list_bus{i,j} = [];
end
end
end
fori = 1:(size(list_bus,1))
iflist_bus{i,8} < 0,0060
for j = 1:(size(list_bus,2))
list_bus{i,j} = [];
end
end
end
fori = 1:(size(list_bus,1))
iflist_bus{i,9} < 0,0053
for j = 1:(size(list_bus,2))
list_bus{i,j} = [];
end
end
end
list_bus( all(cellfun(@isempty,list_bus),2), : ) = [];

fori = 1:(size(list_bus,1))
list_bus{i,1} = datestr(list_bus{i,1});
list_bus{i,5} = datestr(list_bus{i,5},'HH:MM');
list_bus{i,6} = round(86400*mod(list_bus{i,6},1));
list_bus{i,7} = round(86400*mod(list_bus{i,7},1));
list_bus{i,8} = round(86400*mod(list_bus{i,8},1));
list_bus{i,9} = round(86400*mod(list_bus{i,9},1));
list_bus{i,10} = round(86400*mod(list_bus{i,10},1));
end

%generate excel datasheet with output
xlswrite('Bus_Data412',list_bus);
```

A.2 Support functions

Shown below the support functions of the main body sorted in alphabetical order. All of them were own developed, except function *Classify.m.*

A.2.1 Function "*Ali.m*"

```
function [CS3Alis] = Ali(CS3)

fori = 1:(size(CS3,1))
for j = 1:(size(CS3{i,1},1))
for k = 1:(size(CS3{i,1}{j,1},1))
for l = 1:(size(CS3{i,1}{j,1}{k,1},1))
CS3Alis{i,1}{j,1}{k,1}{l,1} = CS3{i,1}{j,1}{k,1}{l,2};
CS3Alis{i,1}{j,1}{k,1}{l,2} = CS3{i,1}{j,1}{k,1}{l,3};
CS3Alis{i,1}{j,1}{k,1}{l,3} = CS3{i,1}{j,1}{k,1}{l,4};
CS3Alis{i,1}{j,1}{k,1}{l,4} = CS3{i,1}{j,1}{k,1}{l,9};
CS3Alis{i,1}{j,1}{k,1}{l,5} = CS3{i,1}{j,1}{k,1}{l,11};
CS3Alis{i,1}{j,1}{k,1}{l,6} = CS3{i,1}{j,1}{k,1}{l,13};
CS3Alis{i,1}{j,1}{k,1}{l,7} = 'off';
CS3Alis{i,1}{j,1}{k,1}{l,8} = CS3{i,1}{j,1}{k,1}{l,5};
end
end
end
end

end
```

A.2.2 Function "*Boa.m*"

```
function [CS3Boas] = Boa(CS3)

fori = 1:(size(CS3,1))
for j = 1:(size(CS3{i,1},1))
for k = 1:(size(CS3{i,1}{j,1},1))
for l = 1:(size(CS3{i,1}{j,1}{k,1},1))
CS3Boas{i,1}{j,1}{k,1}{l,1} = CS3{i,1}{j,1}{k,1}{l,2};
CS3Boas{i,1}{j,1}{k,1}{l,2} = CS3{i,1}{j,1}{k,1}{l,3};
CS3Boas{i,1}{j,1}{k,1}{l,3} = CS3{i,1}{j,1}{k,1}{l,4};
CS3Boas{i,1}{j,1}{k,1}{l,4} = CS3{i,1}{j,1}{k,1}{l,8};
CS3Boas{i,1}{j,1}{k,1}{l,5} = CS3{i,1}{j,1}{k,1}{l,11};
CS3Boas{i,1}{j,1}{k,1}{l,6} = CS3{i,1}{j,1}{k,1}{l,12};
CS3Boas{i,1}{j,1}{k,1}{l,7} = 'on';
CS3Boas{i,1}{j,1}{k,1}{l,8} = CS3{i,1}{j,1}{k,1}{l,5};
end
end
end
```

```
end

end
```

A.2.3 Function "*calculations.m*"

```
function [A] = calculations(B)

a = strncmpi(B{1,3},'Inbound',3);
fori = 1:(size(B,1))
aux = strncmpi(B{i,5},'UQ Lakes',8);
aux2 = strncmpi(B{i,5},'Cultural Centre',8);
aux3 = strncmpi(B{i,5},'Adelaide St',8);
if a == 1
if aux == 1
A{1,12} = B{i,4};
elseif aux2 == 1
A{1,13} = B{i,4};

elseif aux3 == 1
A{1,14} = B{i,4};
end

else
if aux3 == 1
A{1,12} = B{i,4};
elseif aux2 == 1
A{1,13} = B{i,4};

elseif aux == 1
A{1,14} = B{i,4};
end
end
end
```

A.2.4 Function "*Classify.m*"

Courtesy of PhD Candidate in UQ Transport Engineering Mr.Zheng Liang Ma.

```
function R=Classify(CC,CCSize,Raw)

% Classify records according to different criteria
% CC: Classify criteria (a column, such as put the same date record
together)
% CCSize: Size of CC
% Raw: Records to be classified
```

```
% R: Classified results

isstring=iscellstr(CC);
ifisstring==0
    CC=cell2mat(CC);
end

[~,~,xidx] = unique(CC);
Xidx = accumarray(xidx,1:CCSize,[],@(x){x});

for index = 1:size(unique(xidx),1)
Xidx{index} = sort(Xidx{index});
end

R = cellfun(@(x){Raw(x,:)},Xidx);
```

A.2.5 Function "*createlist.m*"

```
function [A] = createlist(A,B)

a = size (A,1);

if B{1,2}==109
A{a+1,1}=B{1,1};
A{a+1,2}=B{1,2};
A{a+1,3}=B{1,3};
A{a+1,4}=B{1,5};
A{a+1,5}=B{1,4};
A{a+1,6}=B{2,12};
A{a+1,7}=B{2,13};
A{a+1,8}=B{2,14};
A{a+1,9}=B{1,17};
A{a+1,10}=B{1,16};
A{a+1,11}=B{1,19};
A{a+1,12}=B{2,19};
A{a+1,13}=B{3,19};

else
A{a+1,1}=B{1,1};
A{a+1,2}=B{1,2};
A{a+1,3}=B{1,3};
A{a+1,4}=B{1,5};
A{a+1,5}=B{1,4};
A{a+1,6}=B{2,12};
A{a+1,7}=B{1,16};
A{a+1,8}=B{1,15};
A{a+1,9}=B{1,18};
A{a+1,10}=B{2,18};
A{a+1,11}=B{3,18};
```

```
End
```

A.2.5 Function "*list.m*"

```
function [A] = list(AUX)

for m = 1:(size(AUX,1))
                SIZE = size(AUX{m,1});
x =AUX{m,1}(:,7);
                [uniqueValues,~,uniqueIndex] = unique(x,'stable');
frequency = (accumarray(uniqueIndex(:),1)./numel(x))*SIZE;
if size(frequency,1)==2
bool=strcmp(AUX{m,1}{1,7},'on');
ifbool==1;
A{m,7} = frequency(1);
A{m,9} = frequency(2);
else
A{m,7} = frequency(2);
A{m,9} = frequency(1);
end
else
                    CMP = strcmp(AUX{m,1}{1,7},'on');
if CMP == 1
A{m,7} = frequency(1);
A{m,9} = 0;
else
A{m,7} = 0;
A{m,9} = frequency(1);
end
end

A{m,1} = AUX{m,1}{1,1};
A{m,2} = AUX{m,1}{1,2};
A{m,3} = AUX{m,1}{1,3};
A{m,4} = min(cell2mat(AUX{m,1}(:,4)));
A{m,5} = AUX{m,1}{1,6};
A{m,6} = max(cell2mat(AUX{m,1}(:,4)));
A{m,8} = 'on';
A{m,10} = 'off';
A{m,11} = AUX{m,1}{1,8};

End
```

B. Statistics

B.1 Descriptive statistics obtained with Minitab

Results for: 412Inbound

Descriptive Statistics: TT; Runn; Dwell; No Passengers

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
TT	226	0	1160,7	9,13	137,2	798,0	1060,0	1145,5	1240,3	1504,0
Runn	226	0	939,8	12,9	193,4	162,0	879,0	960,0	1048,5	1378,0
Dwell	226	0	220,9	11,8	176,9	24,0	105,0	169,5	267,5	902,0
No Passengers	226	0	31,95	1,05	15,81	2,00	20,00	31,00	43,00	68,00

Figure B.1: Descriptive statistics for 412 inbound service

Results for: 109Inbound

Descriptive Statistics: Sect1; Sect2; Total TT; Running _Tim; Dwell _Time; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Sect1	659	36	830,98	6,08	156,11	524,00	738,00	820,00	902,00	3118,00
Sect2	648	47	166,26	6,65	169,25	69,00	118,00	142,00	167,00	2297,00
Total TT	679	16	995,58	9,50	247,64	658,00	872,00	963,00	1063,00	3239,00
Running _Time	679	16	771,84	7,88	205,29	82,00	686,00	753,00	821,00	2850,00
Dwell _Time	696	0	239,3	12,9	340,2	4,0	122,3	201,5	277,0	6877,0
Trip_Demand	696	0	35,487	0,817	21,555	2,000	18,000	31,500	52,000	104,000
Max _passengers	696	0	31,657	0,750	19,780	0,000	15,000	28,000	46,000	95,000

Figure B.2: Descriptive statistics for 109 inbound service

Results for: 412Outbound

Descriptive Statistics: Total TT; Runn Time; Dwell Time; Nº Passengers

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median
Total TT	203	0	1173,5	10,7	152,4	819,0	1059,0	1158,0
Runn Time	203	0	922,25	9,06	129,04	624,00	844,00	892,00
Dwell Time	203	0	251,29	9,01	128,41	0,00	158,00	236,00
Nº Passengers	203	0	28,65	1,03	14,64	2,00	18,00	27,00

Variable	Q3	Maximum
Total TT	1294,0	1499,0
Runn Time	981,00	1418,00
Dwell Time	313,00	658,00
Nº Passengers	38,00	65,00

Figure B.3: Descriptive statistics for 412 onbound service