Module 6 – Case for Analysis Instructions.

1. Perform a Principal Component Analysis (PCA)

a. Results

Ranking 1 for dimension 1: CULTURAL HERITAGE AND EVENTS
Ranking for dimension 2: NATURAL RESOURCES AND SUSTAINABILITY

Region	Ranking 1
Metropolitana	1.838257
Valparaíso	1.218104
Los Lagos	0.109248
O'Higgins	0.064466
Biobío	0.040143
Coquimbo	-0.011710
Araucanía	-0.033997
Maule	-0.117790
Tarapacá	-0.294748
Antofagasta	-0.299452
Arica y Parinacota	-0.359919
Los Ríos	-0.466223
Atacama	-0.487630
Magallanes y Antártica	-0.578266
Aysén	-0.620483

Region	Ranking 2
Valparaíso	0.963874
Los Lagos	0.491080
Coquimbo	0.429749
Metropolitana	0.344312
Los Ríos	-0.025394
Maule	-0.031692
Araucanía	-0.040727
O'Higgins	-0.060718
Arica y Parinacota	-0.090709
Tarapacá	-0.169546
Biobío	-0.309507
Antofagasta	-0.332098
Aysén	-0.335654
Magallanes y Antártica	-0.372268
Atacama	-0.460700

Ranking for dimension 3: HUMAN RESOURCES AND TOURISM-RELATED WORKFORCE DEVELOPMENT
Ranking for dimension 4: TOURISM INFRASTRUCTURE
Ranking for dimension 5: TOURISM MOBILITY AND TRANSPORTATION INFRASTRUCTURE

Region	Ranking 3
Metropolitana	0.972581
Los Lagos	0.355405
Biobío	0.261227
Valparaíso	0.259887
Araucanía	0.072945
O'Higgins	0.062117
Maule	0.030235
Los Ríos	-0.025425
Coquimbo	-0.090273
Tarapacá	-0.178861
Antofagasta	-0.212632
Arica y Parinacota	-0.235170
Atacama	-0.285238
Aysén	-0.395361
Magallanes y Antártica	-0.591436

Region	Ranking 4
Metropolitana	0.936312
Valparaíso	0.577585
Coquimbo	0.315688
Los Lagos	0.280373
Araucanía	0.160874
Biobío	-0.010856
O'Higgins	-0.119660
Los Ríos	-0.121883
Maule	-0.128103
Antofagasta	-0.198182
Arica y Parinacota	-0.226643
Tarapacá	-0.230963
Atacama	-0.296345
Magallanes y Antártica	-0.459335
Aysén	-0.478864

Region	Ranking 5
Metropolitana	1.000101
Los Lagos	0.534727
Valparaíso	0.503329
Biobío	0.259670
Araucanía	0.004996
Arica y Parinacota	-0.030850
Coquimbo	-0.071282
Maule	-0.117912
O'Higgins	-0.139925
Tarapacá	-0.191872
Los Ríos	-0.215598
Antofagasta	-0.235941
Magallanes y Antártica	-0.277069
Atacama	-0.435645
Aysén	-0.586728

b. Steps to apply PCA
    1. Clean the data
        a. Remove symbol $ from income
        b. Remove commas from numbers
        c. Replace empty value with NANs
    2. Deal with the missing value
    3. Apply PCA to the dataset
       Calculate eigenvalues and eigenvectors
       Run PCA and fit the model, choose to use first 7 components
       Calculate the factore scroes

```
Eigenvalues
[ 2.69524913e+01+0.00000000e+00j  1.11294373e+01+0.00000000e+00j
  9.96771135e+00+0.00000000e+00j  9.63014396e+00+0.00000000e+00j
  6.68444451e+00+0.00000000e+00j  4.31780140e+00+0.00000000e+00j
  4.23627924e+00+0.00000000e+00j  3.19391854e+00+0.00000000e+00j
  2.65763895e+00+0.00000000e+00j  2.44088878e+00+0.00000000e+00j
  1.56076049e+00+0.00000000e+00j  1.03352756e+00+0.00000000e+00j
  1.22559809e+00+0.00000000e+00j  1.75507281e+00+0.00000000e+00j
 -3.06139791e-15+0.00000000e+00j -1.85308390e-16+2.20424415e-15j
 -1.85308390e-16-2.20424415e-15j  2.04163097e-15+1.68882770e-16j
  2.04163097e-15-1.68882770e-16j -1.00485326e-15+1.67476022e-15j
 -1.00485326e-15-1.67476022e-15j -1.75964899e-15+6.63798042e-16j
 -1.75964899e-15-6.63798042e-16j  7.44682732e-16+1.60105584e-15j
  7.44682732e-16-1.60105584e-15j  1.55681108e-15+7.46675584e-16j
  1.55681108e-15-7.46675584e-16j -1.63627497e-15+0.00000000e+00j
 -5.01234411e-16+1.48261668e-15j -5.01234411e-16-1.48261668e-15j
  1.56012014e-15+0.00000000e+00j  9.39660698e-16+1.09444618e-15j
  9.39660698e-16-1.09444618e-15j  4.04623807e-16+9.69983123e-16j
  4.04623807e-16-9.69983123e-16j  1.05654050e-15+0.00000000e+00j
 -7.75315260e-16+5.71255638e-16j -7.75315260e-16-5.71255638e-16j
 -1.03844247e-15+0.00000000e+00j -9.29227697e-16+1.69032488e-16j
 -9.29227697e-16-1.69032488e-16j  9.27399350e-16+0.00000000e+00j
  8.94558427e-16+0.00000000e+00j -1.82815705e-16+7.32826510e-16j
 -1.82815705e-16-7.32826510e-16j  7.31948402e-16+3.50640338e-16j
  7.31948402e-16-3.50640338e-16j -4.80439948e-16+4.93617991e-16j
 -4.80439948e-16-4.93617991e-16j  5.97478206e-16+3.57880825e-16j
  5.97478206e-16-3.57880825e-16j -6.00796161e-16+2.03810016e-16j
 -6.00796161e-16-2.03810016e-16j -1.86085719e-16+5.68155159e-16j
 -1.86085719e-16-5.68155159e-16j -4.55622876e-16+3.39780915e-16j
 -4.55622876e-16-3.39780915e-16j  5.83655031e-16+0.00000000e+00j
  2.86180187e-16+4.37512822e-16j  2.86180187e-16-4.37512822e-16j
  4.31565227e-16+3.34628809e-16j  4.31565227e-16-3.34628809e-16j
 -5.28785852e-16+0.00000000e+00j  2.54637145e-16+3.69551356e-16j
  2.54637145e-16-3.69551356e-16j -3.00376387e-16+3.03648175e-16j
 -3.00376387e-16-3.03648175e-16j  4.40483588e-16+1.69217155e-16j
  4.40483588e-16-1.69217155e-16j  4.27998492e-16+0.00000000e+00j
 -4.79357488e-17+2.97800912e-16j -4.79357488e-17-2.97800912e-16j
  2.07034767e-16+0.00000000e+00j  1.41324079e-16+2.03762723e-16j
  1.41324079e-16-2.03762723e-16j -2.37512699e-16+4.56785942e-17j
 -2.37512699e-16-4.56785942e-17j -4.61101194e-17+1.29882423e-16j
 -4.61101194e-17-1.29882423e-16j  1.15885038e-16+0.00000000e+00j
  4.29157872e-17+0.00000000e+00j]
```

```
Eigenvectors
[[-0.00314115+0.j          0.20812988+0.j         -0.06763933+0.j
  ... -0.13568344+0.0505711j  -0.01237383+0.j
   0.1521921 +0.j        ]
 [ 0.18406862+0.j          0.02003813+0.j          0.00677072+0.j
  ...  0.07031605+0.00456505j -0.03526809+0.j
  -0.04038691+0.j        ]
 [ 0.03833594+0.j          0.13057318+0.j          0.19792629+0.j
  ...  0.08731216-0.01717942j -0.10165773+0.j
  -0.01599396+0.j        ]
 ...
 [ 0.15577127+0.j         -0.13419411+0.j          0.07332175+0.j
  ...  0.15025328+0.0551873j  -0.02507885+0.j
  -0.06648522+0.j        ]
 [ 0.0519587 +0.j          0.07674913+0.j         -0.05555667+0.j
  ... -0.02041398+0.12392104j -0.02952919+0.j
   0.03886123+0.j        ]
 [-0.0907754 +0.j         -0.16076641+0.j          0.08952625+0.j
  ...  0.07569427-0.11082345j  0.16648893+0.j
  -0.17456541+0.j        ]]
```

c. Develop a score system for 5 dimension

1. Calculate a weighted average for each variable in principal components (first five rows shown below).

0	1	2	3	4	5	6	7	weighted_average
CULTURAL EVENTS SCHEDULED THROUGHOUT THE YEAR	-0.000975527	0.0266907	-0.00776867	-0.00736376	-0.00998854	0.00791649	0.00662895	0.018019
NUMBER OF CULTURAL CENTERS	0.057165	0.0025697	0.000777646	0.00662925	0.00367884	-0.00227677	-0.00386749	0.076976
WORLD CULTURAL HERITAGE SITES	0.0119058	0.0167448	0.0227327	-0.00146141	0.00432993	0.00578716	-0.00309997	0.067767
NUMBER OF ARCHEOLOGICAL SITES	-0.00224722	1.72511e-05	0.0154391	-0.0177918	-0.0048666	0.0131398	-0.00790565	-0.005017
NATIONAL MONUMENTS	0.0588064	-0.00683228	0.000419941	0.00348383	-0.00428597	7.14347e-05	-0.000379001	0.061038

2. Ranking for each dimension(shown in (a))
3. Identify strengths and opportunity areas for each region, based on the PCA performed.

Combined with the ranking results for each dimension, I could get dimension ranking and average ranking for each region (order by average ranking).

Region	Ranking 1	Ranking 2	Ranking 3	Ranking 4	Ranking 5	avg_ranking
Metropolitana	3.676514	0.688624	1.945161	1.872623	2.000202	1.842620
Valparaíso	2.436207	1.927748	0.519774	1.155170	1.006659	1.278100
Los Lagos	0.218495	0.982160	0.710810	0.560746	1.069454	0.634954
Coquimbo	-0.023419	0.859497	-0.180547	0.631377	-0.142563	0.179016
Araucanía	-0.067993	-0.081454	0.145889	0.321748	0.009993	0.079275
Biobío	0.080286	-0.619014	0.522453	-0.021713	0.519340	0.056952
O'Higgins	0.128932	-0.121437	0.124234	-0.239319	-0.279849	-0.102082
Maule	-0.235580	-0.063384	0.060470	-0.256205	-0.235825	-0.167911
Arica y Parinacota	-0.719837	-0.181418	-0.470341	-0.453285	-0.061701	-0.329883
Los Ríos	-0.932446	-0.050788	-0.050849	-0.243766	-0.431197	-0.354123
Tarapacá	-0.589496	-0.339092	-0.357723	-0.461925	-0.383743	-0.386178
Antofagasta	-0.598903	-0.664197	-0.425264	-0.396363	-0.471883	-0.428825
Atacama	-0.975261	-0.921400	-0.570476	-0.592689	-0.871290	-0.709802
Magallanes y Antártica	-1.156532	-0.744537	-1.182872	-0.918671	-0.554138	-0.753621
Aysén	-1.240967	-0.671308	-0.790721	-0.957728	-1.173457	-0.838493

And then I selected 6 areas to identify their strengths and opportunities (shown in the table below).

Region	Culture	Natural sustainability	Human resource	Tourism infrastructure	transportation infrastructure
Metropolitana	✓	★	✓	✓	✓
Valparaíso	✓	✓	★	✓	✓
Los Lagos	★	✓	★	★	✓
Biobío	★	★	✓	★	✓
Aysen	✓	✓	✓	✓	✓

Team analysis:
- Metropolitana, Valparaíso, and Los Lagos; great example
- Aysén; natural resources and cultural events; but does not have too many HR and infrastructure
- Biobío; cultural events;
- Natural resources and infrastructure
Already doing good in safety and
Improve air transportation and ground transportation

Q&A

1. Project duration & general expectation?
2. Definition to success, increasement of tourists traffic, revenue, ranking in PCA model.
3. environmental sustainability & tourism industry at the same time?

3. In order to improve tourism competitiveness nationwide, I would recommend natural sustainability and transportation infrastructure as the two dimensions to be considered for further investment.
As it is mentioned in OECD, contactless payments and ticketing options in various transportation could save time and money, and kindly be a much safer way which prevent passengers from touching surfaces. More technologically advanced infrastructure can improve quality and decrease cost across the infrastructure asset lifecycle. Also, work from home pattern has been largely applied in many companies. So the investment in the latest information and communication

technologies could greatly increase worker mobility and speed up economic development.

Taken WEF as reference, we can learn that environmentally sustainable tourism plays an important role in the tourism life cycle. The tourism industry must take a leading role in protecting and enhancing the environment on which it depends, for the benefit of the future.