



Developing the TURINFO Model

A Cook-Book Approach



TURINFO - O livro de cozinha

A cook-book to develop the TURINFO Model

A word from the author.

This document is not to be included in the official documentation of TURINFO. It is of internal use and should be regarded as a contribution to training future operators and analyst of TURINFO or something coming after this. It is felt that not enough time had been dedicated to the training of staff members of EMBRATUR and the state tourist offices. The reasons for this were several: Lack of sufficient time, lack of interest of the concerned staff members, physical dislocation, missing preconditions of staff members, lack of priority given to TURINFO by EMBRATUR and the state tourist offices. This should not be understood as criticism of collaboration because the author is not able to judge fully the reasons for each difficulty and shortcomings. Otherwise many experiences of dedicated collaboration made the experience developing something like TURINFO thrilling and rewarding.

Also I am aware of my shortcomings in communication and inability to explain things which seemed clear to me but not to others. This document shall serve as an excuse to those I did not spend the necessary time with, if they would have thought the time was not enough and as a helpful hand to those I never was able to communicate to what TURINFO is all about, although I am aware of the risk to produce just another failing attempt to explain TURINFO.

Many thanks to those who made my work a sea of joy and my comprehension to those who did not.

All these are considerations for further discussion, they are strictly personal and represent in no way the opinion of the Commission of the European Union

Klaus Röder

Contents

1 .Preparation	4
2 .Methodology	5
3 .Data Collection	6
4 .Completing the Data Set	9
5 .Analyzing the Data Set (Factorial Analysis)	15
6 .Elaborating the System of Equations	16
7 .Filling the Model Spreadsheets	22
8 .Creating the interactive information system	27
9 .Documenting the model and its results	29
10.Making TURINFO work	31

1 Preparation

The purpose to develop an econometric model for the Brazilian States was to show the contribution of tourism as an economic factor. This required some preliminary considerations. The question was:

- Who will want this model?
- Who will use it?
- Which will be the technical platform to use it?
- What should be the benefit of the model?
- Which are the data available to elaborate the model?
- Which are the technical necessities to elaborate the model?

To some of the questions no reliable answer was possible, especially not from a technical consultant who is not familiar with the general conditions of the country involved and the special circumstances of the organization the project management has to cooperate with, since he was not involved in the project definition and preparation.

- The model was developed for the benefit of the State Tourist Organizations. They should be enabled to use this model in their effort to promote the tourist trade in the different regions and states.
- The persons to use it would be in the first place the staff members of these organizations, whereas the persons to be informed about the subjects dealt with in the model, would be decision makers, entrepreneurs in the tourist business in the first place, economists, statisticians, people interested in tourism, econometry, statistics, modelling etc. in the second place.
- The technical platform should be as lean as possible without sacrificing the possibilities of advanced micro-computer technology.
- The benefit of the model should be simple and interactive use of its facilities. The model user should be enabled to see past influences and reasons of the development of economical indicators. The model should enable the user to project future developments of economical events in a reasonable time span.
- The data available for the elaboration of the model proved to be rather inhomogeneous for the different states. There have been economical censuses however in an interval of 5 years in the past, which can be considered a reliable and consistent data base. Unfortunately the 1990 census was denied financing by the Collor-Government. The next census however is scheduled for 1995 and should considerably improve the data

situation as far as economical data are concerned. The situation about tourism data is different from state to state. Whereas the data about international tourists are rather homogeneous (EMBRATUR conducts a regular survey three times a year, countrywide since 1985) the data about Brazilian tourists in the states depend on the efforts of the state tourist's organization. Some states have never conducted any survey (like **Rio de Janeiro**, but there was once a survey about tourism in the city), other have just started (**Distrito Federal**), others have conducted surveys for some years to find out about the demand of national tourists within the state (**Parana, Bahia, Amazonas**).

- In view of these preconditions which were not all known to the technical consultant at the time the project started, the decision about the technical necessities were the following:
 - Use of a spreadsheet as the main software support for data entry and development of the model with the maximum of presentation capabilities (**Quattro Pro** 1.0 and 5.0 for Windows).
 - Use of a statistical package like SAS or SPSS for statistical analysis (**SAS** and **SPSS**)
 - Use of a powerful word processor for system documentation (**WordPerfect** 6.0 for Windows and DOS for documentation in general, **Corel Draw** for Windows for Illustrations and Title Pages)

2 Methodologies

The methodology adopted is to use a set of straightforward multiple regression equations (the number is about 60) with tourism demand as the set of independent variables and the economic data available as the set of dependent variables. Dummy variables will enter the equations if necessary but usually the development of economic data will be expressed as being accompanied (or influenced) by tourist demand only.

Economic data of tourism are the data of 14 activities considered as related to tourism of the services sector, being three of the commerce sector, 8 of the other services sector (or 7 of commerce and 4 of other services if hotels and restaurants are included into commerce) and three of transports.

Tourist demand was expressed among other information as tourist arrivals (Brazilians and foreigner), tourist spending in the state (Brazilians and foreigner) and tourist spending in the hotels of the state (Brazilians and foreigner).

Time series available over the passed years (1980 until 1992) will enter the equations to observe development of dependent and independent variables in the past and allow projections for the next (three) years assuming a standard behaviour of the independent variables of tourist demand. For projection purposes a standard increase in 5% of the variables of tourist demand was assumed. This is not an assumption of future behaviour of tourist demand but an assumption which can be modified in various case studies.

This technique is used in all of the 11 model spreadsheets except two. A percentage of tourist revenues compared to Gross Domestic Product values in four graphics is calculated in one of the spreadsheets (*UF_PIB02.WB1*)¹. Since there are no equations in this spreadsheet, the future values are calculated with moving averages, usually using the last 4 years to be included in the calculation. The second spreadsheet producing model results includes straightforward series of tourist demand (*UF_TUR01.WB1*). In this spreadsheet, the data series of tourist demand are displayed in 10 graphics. Sometimes percentage comparisons are calculated, sometimes totals are calculated out of average values entered from the time series.

Before series enter in equations the whole data set is analyzed with Factorial Analysis for two reasons:

- Discover coherence within subgroups of the various variables
- Allow to chose independent variables in multiple regression and avoid auto-correlation.

More discussion about the techniques in further chapters. For more discussion about why no other methodology was chosen, see *TURINFO: Introdução ao sistema*

3 Data Collection

Data are mainly of four different origins:

- The data of the economical census conducted by the IBGE (Instituto Brasileiro de Geographia e Estatística) in 1980 and 1985
- Macroeconomic data, like GDP of the state, population etc.
- Tourist demand as tourist arrivals (Brazilians and foreigner), tourist spending in the state (Brazilians and foreigner) and tourist spending in the hotels of the

¹ *UF* stands for Unidade Federal, it is replaced by the two letters representing the Brazilian State, e.g. **PR** for Parana or **BA** for Bahia.

state (Brazilians and foreigner) mentioned before, but also data like means of spending for hotel and means of total spending and number of days spent as an average by foreign and Brazilian tourists.

- The census data of 1980 and 1985 **aggregated** for small, medium and big companies bought from IBGE and delivered on diskette in 12 text files bought from project funds in April 1994.

These data are entered in 8 spreadsheets of QUATTRO PRO. The names of these spreadsheets are the following:

For the first data set:

- *UF_COM80.WB1* for entering the data of the commercial census in 1980.
UF_COM85.WB1 for entering the data of the commercial census in 1985.
UF_SER80.WB1 for entering the data of the other services census in 1980.
UF_SER85.WB1 for entering the data of the other services census in 1985.
UF_TRA80.WB1 for entering the data of the transports census in 1980
UF_TRA85.WB1 for entering the data of the transports census in 1985

In these spreadsheets the data are copied from the IBGE printouts which EMBRATUR bought from the IBGE. How this is done is described in detail in *TURINFO:Introduction to the system -Chapter 4 and Annex E.* , the names of the spreadsheets have changed slightly, but the way of using them has remained.

For the second data set:

- *UF_PIB_E.WB1* for entering the macroeconomic data. These data are mainly the GDP of the state and the GDPs of the three different subgroups: commerce, other services and transports as well as some other general data. The data should be entered straight away as time series from 1980 to the most actual year. As far as financial data are concerned, they should be given in US Dollars at current prices. A column with official conversion rates of the Brazilian Central Bank is included in these spreadsheets and Cruzeiro values should be converted to US\$ dividing each Cruzeiro value by its corresponding conversion rate. For the time being values are in current or nominal values instead of real values. It would be easy to transform all financial values with a Dollar-deflator into real values, but the importance of the system is **not** to show real growth in production, but the effect of tourism demand on economic development. The multipliers or regression coefficients gained as a result of this system are probably overestimated because of this

approach, but what is more important is to show the interaction between tourism and economy and this is also plausible with current prices. Since curves are usually more accentuated, influence and impact can be shown and demonstrated more clearly.

For the third data set:

- *UF_TUR_E.WB1* for entering the tourism data. These data are mainly the data collected by EMBRATUR about foreign tourists and if available data from the state tourism organizations about Brazilian tourist demand. Like all spreadsheets the columns have short variable names which are explained in the accompanying documents. Again financial data should be collected in US Dollars at current prices.

For the fourth data set:

- *com80gm.dat* contains data about medium and big establishments: census commerce 1980
com80pg.dat contains data about small enterprises: census commerce 1980
com85gm.dat contains data about medium and big enterprises: census commerce 1985
com85pg.dat contains data about small enterprises: census commerce 1985
- *ser80gm.dat* contains data about medium and big enterprises: census other services 1980
ser80pg.dat contains data about small enterprises: census other services 1980
ser85gm.dat contains data about medium and big enterprises: census other services 1985
ser85pg.dat contains data about small enterprises: census other services 1985
- **(*trp80gm.dat* contains data about medium and big enterprises: census transports 1980
trp80pg.dat contains data about small enterprises: census transports 1980
These two data sets in parentheses were actually missing on the diskettes and should be delivered as soon as possible by IBGE)
trp85gm.dat contains data about medium and big enterprises: census transports 1985
trp85pg.dat contains data about small enterprises: census transports 1985**

The original data used were the ones from the first set including only medium and big enterprises (more than 5 employees and a cut-off revenue). This excluded however many enterprises in certain branches, e.g. travel agencies or boarding houses. This then lead to distorted number as far as number of enterprises were concerned. So the decision was to include also data of small enterprises into the model. These 12 data sets mentioned above were used to complete the original data sets as described in the next chapter.

4 Completing the Data Set

Data spreadsheets mentioned above should be stored in subdirectory belonging to the states data, e.g. *UF_BASE*². Then load spreadsheet *MACLIB03.WB1*. This macro library is for all kind of miscellaneous macros (e.g. ALT+<T> fixes title lines and columns at cursor position) but mainly for completing data sets into data tables. Macro names are listed at column A:A1, addresses in A:B1, description in A:C1. Use the spreadsheet *MACLIB03.WB1* of the subdirectory involved for the development of the model or create new subdirectories with the appropriate names. If data structure varies, e.g. hotels and restaurants are included into commerce (as in Distrito Federal and Bahia) the macro library also varies.

UF stands for the *n*th-states abbreviation, e.g. DF for Distrito Federal. The first and main directory *UF_BASE* contains spreadsheet files and macro file *MACLIB04.WB1*, *UF_BASEB* contains backup files and *UF_DOCU* contains files necessary for the documentation, see chapter 9.

MACLIB03.WB1 of a comparative state should be copied to the appropriate subdirectory *UF_DOC*, it contains a macro library which is briefly documented in the B column, addresses of macros are given in column A.

The main macros to use for data set completion:

macro:{C80S} make first row of TABLE1 (commerce) out of *UF_COM80.WB1*

macro:{C85S} make fifth row of TABLE1 (commerce) out of *UF_COM85.WB1*

macro:{S80S} complete first row of TABLE1 (services) out of *UF_SER80.WB1*

² The subdirectory structure for the TURINFO data is the following, as least on the project design and the project computer:

```

c:      /-dbases /-uf1_base   /-uf1_baseb
          /-uf1_docu
          /-uf2_base   /-uf2_baseb
          /-uf2_docu
          /-ufn_base   /-ufn_baseb
          /-ufn_docu

```

macro:{S85S} complete fifth row of TABLE1 (services) out of UF_SER85.WB1
 macro:{T80S} complete first row of TABLE1 (transports) out of UF_TRA80.WB1
 macro:{T85S} complete fifth row of TABLE1 (transports) out of UF_TRA85.WB1

All of these macros need the indicated spreadsheet to be open and a spreadsheet with the name of MATMOD01.WB1 to be opened, which should be empty at the beginning or should hold only year numbers in the first column. Macros are executed in QUATTRO PRO by pressing <ALT>+<F2> and indicating then the macro name, either typing it or choosing it out of a selection-list of macros.

Each of these macros consists of several parts:

- a. cleaning the non-US Dollars currency values and copying the Dollar values in the columns following the variable names. Summing the appropriate values to aggregate into the 8 aggregated variable-names: NES_S\$\$, AQU_S\$\$, PER_S\$\$, SAL_S\$\$, INP_S\$\$, DES_S\$\$, REC_S\$\$ and VAL_S\$\$\$. The first \$ stands for the number of the series, the second \$ stands for the activity group (C for commerce, S for services, T for Transports). The 112 series names (8*3 for commerce, 8*8 for services, 8*3 for transports) are created by the macros.
- b. These variable names are included in new columns in front of each data column (according to the number of the series).
- c. The variable names and the values are transposed and spread according to the format of TABLE1 and copied to spreadsheet MATMOD01.WB1

The spreadsheet MATMOD01.WB1 should look like this after having executed all of the 6 macros (this is only a partial view of the beginning):

ANO_VAR	NES_S1C	AQU_S1C	PER_S1C	SAL_S1C	DES_S1C
1,980	3,981	32,509,515	18,988	29,230,087	418,927,165
1,981					
1,982					
1,983					
1,984					
1,985	4,266	25,743,662	20,987	21,711,342	369,947,730
1,986					
1,987					
1,988					

1,989					
1,990					
1,991					
1,992					

Observation: Use Quattro Pro 5.0 to execute the macros or change all format statement ({\Block;Format...}), they did not translate correctly from Quattro Pro DOS. Any change in the aggregation of the IBGE data should be altered in the macros, if for further processing the print-outs will be used at all. It would be more comfortable to use the disaggregated data from the IBGE on diskette. These have not been provided yet. In the actual models the data for medium and big enterprises have been completed with the data for small enterprises. Unfortunately the data delivered by IBGE on diskette contain only aggregated data. The data for medium and big enterprises are aggregated in about the same way as in the TURINFO model (Number, Employees, Salaries etc.) The only missing values are the investments. The data are available for each of the requested activities. The census data which are missing however are the ones from **TRANSPORT 1980**. The data for small enterprises are aggregated in the same way but not for activities (4 digits, e.g. 5111 for hotels) but for activity groups (3 digits, e.g. 511) comprising three different activities of accommodation services. The data for small enterprises have been distributed proportionally to the data for the medium and big enterprises and added to these. This was due to the observation of Tourism organizations, that data values were 'too low', especially for number of enterprises and employees. Now the data are probably too high due to overestimation and having joined values from activities not included in the 14 'tourism' activities.

AN URGENT REQUEST FOR THE FUTURE: A complete data set of non-aggregated data of the census 1980 and 1985 is needed for all types of enterprises for the 14 'tourism' activities, all the 27 states and all the 6 censuses. These data should be acquired on electronic media. If data are acquired in disaggregated form (as they appear in the questionnaire) future aggregation can be modified and controlled.

Completing the data set, means to fill the missing values for the years between and after the census data. Again three macros in MACLIB03.WB1 are used to facilitate the completion of spreadsheet MATMOD01.WB1.

The basic idea is to use the values of the two census years as base points, include an arithmetic mean between the two values and add eventually further values (calculated by arithmetic means) to add enough values necessary for regression, i.e. avoid indefinite matrix, as **dependent values**.

As **independent variables** take the corresponding years and at least one correspondence series. This has been in most cases series of the corresponding GDP, because it is thought that evolution of the series will follow roughly that of the GDP of the state.

Observation: This is of course another source of error, the values between the census years are estimated and not observed, although they are treated as observed values.

The name of the spreadsheet is MATMOD02.WB1. The name of the macros (in MACLIB03.WB1) are REG1 for the series belonging to commerce, REG2 for the series belonging to services and REG3 for the series belonging to transports. If the number of series included into commerce etc. changes (e.g. if hotels and restaurants are included) these macros have to be modified. Two varieties appear on the sample diskette (in PR only three activities are included into commerce, in DF there are seven, including hotels and restaurants). The series of reference included are the respective GDPs, but this can vary, if better references can be found for estimation the evolution of the regressed series, e.g. series from transport surveys to serve as a reference to complete the transport series.

Other series of reference have to be substituted in the position of the former series. Still there might be necessity to modify the completed series manually. Physical and financial values may not become negative. This can be the case if the value of a series in 1985 is much inferior of that of 1980.

The result should be a completed matrix of 13 rows (1980 until 1992) of 122 series,(14 activities and 8 series each). The macroeconomic series *UF_PIB_E* and the tourist demand series of *UF_TUR_E.WB1* should be appended column wise to these matrixes. The result is the so called TABELA, is stored under the name of *UF_TAB1.WB1* in the respective subdirectory *UF_DOCU*.

The sample file is called MATMOD1C.WB1. Watch the observed data being maintained in the completed data set::

ANO_VAR	NES_S1C	AQU_S1C	PER_S1C	SAL_S1C	DES_S1C
1,980	3,981	32,509,515	18,988	29,230,087	418,927,165

1,981	4,029	31,359,280	19,328	27,951,856	410,600,374
1,982	4,072	30,336,882	19,630	26,815,687	403,199,018
1,983	4,162	28,217,949	20,256	24,460,963	387,859,617
1,984	4,219	26,843,759	20,662	22,933,856	377,911,567
1,985	4,266	25,743,662	20,987	21,711,342	369,947,730
1,986	4,291	25,148,938	21,163	21,050,438	365,642,396
1,987	4,348	23,801,481	21,561	19,553,038	355,887,869
1,988	4,389	22,806,443	21,855	18,447,274	348,684,579
1,989	4,406	22,404,105	21,974	18,000,164	345,771,967
1,990	4,459	21,152,472	22,343	16,609,252	336,711,136
1,991	4,543	19,158,053	22,932	14,392,897	322,273,112
1,992	4,619	17,355,664	23,465	12,389,942	309,225,240

Regressed data series and reference series:

corr. reg.	lin	reg	pibc for series 1-3	pics for series 4-11	pibt for series 12-14
98,685,068	98,595,126	98,780,097	3,636,688,528	14,892,823,876	1,332,382,853
100,589,242	101,358,869	100,684,271	3,818,541,007	17,289,785,328	1,721,831,062
102,861,008	104,122,612	102,956,036	4,170,910,692	18,089,163,528	1,975,979,005
107,328,715	106,886,355	107,423,744	2,740,287,630	13,381,328,882	1,421,860,329
110,148,426	109,650,098	110,243,454	2,520,635,285	12,957,318,518	1,474,324,749
112,413,842	112,413,842	112,508,870	2,746,664,327	13,319,069,771	1,730,809,679
114,611,936	115,177,585	114,706,964	3,794,442,669	17,146,026,161	2,012,074,046
117,346,933	117,941,328	117,441,961	3,618,258,338	17,657,264,033	2,095,719,038
119,764,735	120,705,071	119,859,763	4,015,116,600	19,732,607,681	2,296,114,881
120,194,413	123,468,815	120,289,442	5,375,719,304	27,214,292,765	3,228,284,913
122,225,012	126,232,558	122,320,041	5,355,348,351	27,462,814,914	3,571,199,432
126,746,549	128,996,301	126,841,578	4,127,186,864	22,937,985,111	2,997,267,740
129,232,043	131,760,044	129,327,072	3,211,272,464	21,830,931,426	3,172,747,956

Each column is regressed consecutively by the macros. Regression results: (dependent variable, column 2) and independent variables (column 3 and 4). These results change during the iteration:

	98,595,126	1,980	1,332,382,853
--	------------	-------	---------------

	105,504,484	1,983	1,698,919,6 67
	112,413,842	1,985	1,730,809,6 79
	119,323,200	1,989	3,571,199,4 32

The three different tables commented above are all part of the spreadsheet MATMOD1C.WB1. The revised resulting matrix is TABLE1.

This matrix is transformed into TABELA2 (*UF_TAB2.WB1*) with macro T1T2 in MACLIB03.WB1 (again watch the version of MACLIB03.WB1, if the number of series to be summed up varies, the macro has also to be different) and into TABELA3 (*UF_TAB3.WB1*) with macro T2T3. The matrixes differ in that respect, that they contain sector sums (TABELA2) and total sums (TABELA3) for the 8 types of collected variables, all of them contain the macro-economic series and the tourist demand series. The matrixes can be saved as LOTUS (*.WK1) files for data conversion by other software products.

5 Analyzing the Data Set (Factorial Analysis)

The first analysis is performed running a factorial analysis on the data set. The examples given are in SPSS/PC+ (V4.0) for two reasons. SPSS/PC+ is less space consuming on the hard disc and allows easy and rapid conversion of LOTUS (*.WK1) files. Otherwise SPSS allows all the necessary analysis routines for the development of TURINFO, as does SAS. The problem to translate the sample SPSS-routines into SAS programs should be a minor problem.

A typical SPSS-program for f.a. looks like this:

```

1 TRANSLATE FROM 'c:\dbases\df_base\df_tab2.wk1'/FIELDNAMES.
2 SAVE OUT="df_fac2.SYS"/DROP ANO_VAR CHTI_VAE TO CHTI_FLU
3 GASTP_ES TO GASHT_ES GASTN_ES GASHN_ES CUSTVI_E INFUN_PA TO
INFUN_EM.
4 GET FILE 'df_fac2.SYS'.
5 factor variables=ALL
6 /print kmo
7 /plot eigen

```

8 /criteria mineigen(1.1)
9 /save reg (4,faktor).
10 list faktor1 to faktor4.
11 factor variables=ALL
12 /print kmo repr
13 /analysis=all
14 /criteria mineigen(1.1)
15 /extraction=pc
16 /rotation varimax
17 /rotation oblimin.

Lines are numbered in order to identify them, numbers are NOT part of the program text. The factor analysis (f.a.) uses

Line 1: Translates Lotus Spreadsheet into SPSS-file format, maintains column headings as variable names

Line 2 and 3.: Drops several variables, estimated not to be important for f.a.

Line 6: Print Kaiser-Meyer-Olkin measure, this should be above 0.8

Line 7: Plot Eigenvalues

Line 8: Criteria to select Eigenvalues

Line 10:List normalized Eigenvalues

Line 16 and 17:Two rotations-methods to ease interpretation of factors: **varimax** which maintains orthogonality of factors and **oblimin** which does not, but eases further interpretation of factor meaning. For further use of results of f.a. see *TURINFO: Dados Técnicos do Modelo relativo...* and bibliography.

Factorial analysis is not essential for elaborating TURINFO but helpful in that respect that it gives an overview over the huge data set and limits the risk of selecting auto-correlated series in the models.

6 Elaborating the System of Equations

The next step in elaborating the system is to develop the set of equations. These are developed using the three matrixes: TABLE1, TABLE2 and TABLE3. Each system uses 13 spreadsheets:

UF_ENT01.WB1 contains no equations, contains initial slide show and selection graphics

<i>UF_PIB01.WB1</i>	contains 4 equations concerning PIB explained by tourist demand
<i>UF_PIB02.WB1</i>	contains no equations concerning participation of tourist demand in PIB
<i>UF_NES01.WB1</i>	contains 4 equations concerning number of tourism establishments explained by tourist demand
<i>UF_AQU01.WB1</i>	contains 9 equations concerning investments of tourism establishments explained by tourist demand
<i>UF_PER01.WB1</i>	contains 9 equations concerning employment in tourism establishments explained by tourist demand
<i>UF_SAL01.WB1</i>	contains 4 equations concerning salaries paid by tourism establishments explained by tourist demand
<i>UF_INP01.WB1</i>	contains 4 equations concerning transactions of tourism establishments explained by tourist demand
<i>UF_DES01.WB1</i>	contains 4 equations concerning operational expenses of tourism establishments explained by tourist demand
<i>UF_REC01.WB1</i>	contains 9 equations concerning operational revenues of tourism establishments explained by tourist demand
<i>UF_VAL01.WB1</i>	contains 9 equations concerning operational surplus (or value added, depending on the system) of tourism establishments explained by tourist demand
<i>UF_TUR01.WB1</i>	contains no equations, displaying series of tourism demand
<i>MACLIB04.WB1</i>	contains the macros to link the system graphics and spreadsheets

This makes a total of 56 equation for each system, if all data were available.

The exact name of the dependent variables might vary slightly, but in the case of 4 equations in the spreadsheet, it is usually the variable for all the tourist establishment, e.g. *PER_TTT* (Total of employees in the tourist activities in the State) and the variable for each of the three different sectors, e.g. *PER_T_C* (Total of employees in the tourist activities of commerce), e.g. *PER_T_S* (Total of employees in the tourist activities of services), e.g. *PER_T_T* (Total of employees in the tourist activities of transport). If 9 equations appear in the spreadsheet, then the series for 5 selected activities, series 1,3 and 8 of services and 1 and 3 of transports are selected as dependent variables too, e.g. *PER_S1S* (Total of employees in the hotel sector), e.g. *PER_S3S* (Total of employees in the restaurants sector), e.g. *PER_S8S* (Total of employees in the travel agencies sector), e.g. *PER_S1T* (Total of

employees in passenger road transport sector), e.g. PER_S3T (Total of employees in passenger air transport sector).

Each of the equations is elaborated with the help of a statistics program, like SPSS or SAS. A typical sample program in SPSS looks like this, again line numbers are not part of the program code:

```
1 TRANSLATE FROM 'c:\databases\df_base\df_tab3.wk1'/FIELDNAMES.
2 SAVE OUT='df_genf3.SYS'/drop ano_var.
3 GET FILE 'df_genf3.SYS'.
4 REGRESSION /VARIABLES
5 nes_ttt
6 num_tint num_tnat
7 gas_tnat to gho_tint
8 /stat ci coeff r
9 /DEPENDENT
10 nes_ttt
11 /method back
12 /res durbin.
```

Line 1: translates spreadsheet (TABLE3 of DF in this case) into SPSS file format.

Line 2 and 3: use and maintain internal file format in order to avoid repeated translations of spreadsheets.

Line 4: command to start regression

Line 5 to 7: variables included in the model, for simplicity this example uses one dependent variable (nes_ttt) and 6 (possible) independent variables, which will be selected by the regression routine.

Line 8: statistics calculated: R , R^2 , corrected R^2 and Standard Error of estimate. F value is printed after each change in model together with significance. Regression coefficients together with Standard error, beta-coefficient and t-values, 95% confidence interval for regression coefficients. (This is only a proposal, other statistics can be calculated)

Line 10: Dependent variable

Line 11: Method selected, *backwards* proved to be quite efficient, in many cases *stepwise* might be better to avoid too many independent

variables. This depends on the actual data set and no final recommendation about the method can be given

Line 12: Coefficient Durbin-Watson

The acceptance of each model is done according to the criteria named in the document:

TURINFO: Introdução ao sistema, page 41.ff:

1. Does the model make sense ?
2. Coefficient R^2 (close to 1) and significance of distribution F (<0.05)
3. Distribution t of regression coefficients and their significance levels (<0.05)
4. Durbin-Watson coefficient measuring possible auto-correlation between variables ($1.5 < dw\text{-coefficient} < 2.5$)

A typical final output would look like this:

Equation Number 1 Dependent Variable.. NES_TTT

Variable(s) Removed on Step Number

11.. GHTI_E6

Multiple R .93511

R Square .87443

Adjusted R Square .86301

Standard Error 1232.05588

F = 76.59795 Signif F = **.0000**

Variable B SE B 95% Confdnce Intrvl B Beta

GATI_E6 **1.312482E-04** 1.49963E-05 9.824152E-05 1.642549E-04 .935108

(Constant) 14474.219976 751.323714 12820.567786 16127.872166

Variable T Sig T

GATI_E6 8.752 **.0000**

(Constant) 19.265 **.0000**

Residuals Statistics:

Min Max Mean Std Dev N

*PRED 16939.3262 26818.6250 20330.3716 3112.7794 13

*RESID -2266.5195 1824.4414 .0000 1179.6037 13

*ZPRED -1.0894 2.0844 .0000 1.0000 13

*ZRESID -1.8396 1.4808 .0000 .9574 13

Total Cases = 13

Durbin-Watson Test = **1.63912**

In this specific case all the demand variables are removed until only one (GATI_E6) remains. All the selection criteria were fulfilled.

According to regression theory the independent variable: Expenses of foreign tourist 'explains' the dependent variable: number of tourist establishments in the state. With each change of one Dollar of expenses the number of establishment varies according to the regression coefficient (1.312482E-04) or in other terms, $1/0.0001312482 = 7620$ Dollars must be spent by foreign tourist to expect one more tourist establishment in the state. Remember that regression never states that one variable 'causes' the change of the other but rather accompanies the development. It's a mathematical relation and the analyst has to 'explain' it.

In several cases variables with negative sign enter the equation. Assuming the fact, that the independent variables are not auto-correlated, which is confirmed by an acceptable DW-coefficient, each regression coefficient can be interpreted independently however (if all the other variables remain unchanged). If these automatic selections and deselection of independent variables seems inconvenient (e.g. the analyst wishes, that expenses of Brazilian tourists are included by all means in the list of independent variables) or if the Durbin-Watson coefficient renders bad results, independent variables have to be selected manually (e.g. with the help of factorial analysis, chose a variable from a different factor to make auto-correlation less likely) or chose a different method (*forward, stepwise, enter etc.*) to select new variables.

The result of the regression analysis is an output file (or several files) with regression coefficients, constants and eventually confidence intervals for the coefficients.

7 .Filling the Model Spreadsheets

The twelve model spreadsheet contain all in the upper left section the independent variable series, the projected data for years 1993 to 1995 are calculated multiplying the values of the previous year with 1.05 , this is an assumed 5% increase, but these assumptions can be modified in scenarios. This part of the spreadsheet named BA_PIB01.WB1 looks like this:

Dependent Variable..	CHTI_ES	CHTTNES	GATI_E6	GATN_E6	GHTI_E6
PIBT_EST					
PIB do estado					
'80	121,000	1,494,000	23,522,058	167,370,906	8,824,697
'81	75,000	1,491,000	18,782,020	179,112,830	6,620,482
'82	85,000	1,508,000	21,487,245	179,659,236	7,976,870
'83	116,000	1,707,000	22,776,374	156,580,573	8,338,279
'84	152,000	1,784,000	30,214,859	163,421,987	11,949,379
'85	171,000	1,888,000	32,345,397	164,999,709	12,763,614
'86	208,000	2,338,000	43,708,554	226,008,143	17,338,931
'87	190,000	2,100,000	43,202,469	220,139,146	16,958,581
'88	185,000	1,855,000	43,490,786	202,365,288	17,205,146
'89	222,000	1,952,000	74,881,628	273,331,491	29,537,144
'90	178,000	1,548,000	63,961,851	262,223,350	23,985,694
91	202,000	1,782,000	67,618,688	226,769,991	28,678,387
92	257,000	1,708,000	94,053,894	282,907,402	41,107,797
(93)	269,850	1,793,400	98,756,589	297,052,772	43,163,187
(94)	283,343	1,883,070	103,694,418	311,905,411	45,321,346
(95)	297,510	1,977,224	108,879,139	327,500,681	47,587,414
Variable PIBT_EST		B 95% Confdnce Intrvl			
CHTTNES	3521.842	1729.1941	5314.4899		
GATI_E6*	274.8182	77.364352	472.27205		

GHTI_E6	-521.1247	-972.3897	-69.85965		
(Con)	1033395616	7126913056	1354099926		
	2		9		

The lower part contains the regression coefficients and the constants as well as the confidence intervals for all coefficients.

The regression formulas enter in a different part of the spreadsheet, usually right of the series of independent variables data. They are shown together with the formulas in the first three cells. The formulas are the regression equations for 1980. they are simply copied down for all the subsequent years. The formulas do not appear like this, but render the results of the table. Watch the fact that the formulas differ only in the constant \$B\$24, \$C\$24, \$D\$24. The different formulas calculate the lines illustrating the confidence intervals.

+ \$A:\$B\$26+\$C3*\$A:\$B\$23 +\$D3*\$A:\$B\$24+\$F3*\$A:\$ B\$25	+ \$A:\$B\$26+\$C3*\$A:\$B\$23 +\$D3*\$A:\$C\$24+\$F3*\$A:\$ B\$25	+ \$A:\$B\$26+\$C3*\$A:\$B\$23 +\$D3*\$A:\$D\$24+\$F3*\$A:\$ B\$25
res	resconlo	resconhi
17.461.110.279	12.816.589.439	22.105.631.141
17.296.566.962	13.587.984.816	21.005.149.127
17.393.036.005	13.150.296.905	21.635.775.128
18.259.819.404	13.762.536.856	22.757.101.975
18.693.399.143	12.727.359.010	24.659.439.307
19.220.863.728	12.834.140.599	25.607.586.889
21.544.184.469	12.913.762.382	30.174.606.599
20.765.114.133	12.234.620.551	29.295.607.758
19.853.006.501	11.265.583.632	28.440.429.414
22.394.891.612	7.609.226.187	37.180.557.111
20.864.111.480	8.234.597.916	33.493.625.109
20.247.710.027	6.896.139.999	33.599.280.122
20.774.697.084	2.203.393.960	39.346.000.303
21.296.734.131	1.796.865.850	40.796.602.510
21.844.873.029	1.370.011.334	42.319.734.827
22.420.418.872	921.814.093	43.919.023.761

In each equation there are usually three columns each, even if the lines for the confidence intervals are not shown in the graphics. The titles are named res0 .. res8 (for the regression results), res0lo ... res8lo (for the low value of confidence interval) and res0hi..res8hi (for the high value of confidence interval) respectively.

The coefficients are preserved in the form as they appear in the statistic programs output files. The resulting ASCII files of the regression analysis of the statistics program should be copied to the spreadsheets.

Observation.: Quattro Pro Vers 5.0 has the better options to import ASCII files, i.e. the 'Parse Expert'.

Exceptions are the spreadsheets

<i>UF_ENT01.WB1</i>	contains no equations, contains initial slide show and selection graphics
<i>UF_PIB02.WB1</i>	contains no equations concerning participation of tourist demand in PIB
<i>UF_TUR01.WB1</i>	contains no equations, displaying series of tourism demand and of course
<i>MACLIB04.WB1</i>	contains the macros to link the system graphics and spreadsheets

These spreadsheets contain no equations, the first three contain graphics however. The graphics are named according to the convention found in the document *TURINFO: Introdução ao sistema. Annex K.*), they contain the dependent series and at least one independent series. In the help screen, the regression coefficients appear as far as they can be explained and interpreted. The above mentioned spreadsheets *UF_PIB02.WB1* and *UF_TUR01.WB1* calculate series with simple arithmetic means, like dividing total revenues by GDP value to show percentages of revenues against GDP. Graphs are created based on these calculated series. The spreadsheets together with their graphs should be easily understood. For technical details about editing graphs consult the bibliography, especially the user's manual of Quattro Pro.

Some spreadsheets contain scenarios, a graph with a dialogue box to simulate future development, i.e. change dependent variable values and watch the

results or indicate desired results and observe necessary independent values to obtain these results.



In all spreadsheets with scenarios this graph is called 'Gráfico3' in the list of graphs. It is usually located as 'floating graph' on page B of the notebook. QUATTRO PRO calls the collection of

256 spreadsheets under one name a notebook.

The example shows this graph together with the 'Speed buttons' to operate the scenario. The first button shows the indicated dialogue box, the second calls the 'solve for' dialogue box, the third returns to the calling spreadsheet, the fourth shows the help screen, the fifth prints the actual graph on the standard Windows Printer (if one is attached).

This example of programming a scenario might illustrate the use of macros in Quattro Pro. For more details see *QUATTRO PRO, Construindo Aplicativos com Planilhas*.

The first button: calls Macro {DoDialog *Series1,A25*}. The part in italics is optional, the first item is the name of the dialog box, which appears like this in the graphics page. The second is a cell in the actual spreadsheet which has to be indicated but is not further used. The dialog box must be designed before with the menu command Dialog Links to link each object in the dialog box with the corresponding spreadsheet data, e.g. link the rollbars with data cells in the spreadsheet, which are directly shown in 'Gráfico3'. There are mainly 3 types of objects: 1 the rollbars (reflect the data series of the independent variable for the respective year), the edit field below (shows the value of the independent variable for the respective year) and the OK and CANCEL button (closes dialog box). The graph simply reflects the values of dependent and independent series.

The second button: calls Macro {DoDialog *Meta1,A67*}. The part in italics is optional, the first item is the name of the dialog box, which appears like this in the graphics page. The second is a cell in the actual spreadsheet which has to be indicated but is not further used. The dialog box must be designed before, the menu command Dialog Links for each object in the dialog box have to be established. There are mainly 2 types of objects: The edit fields (shows the value of the dependent variable for the respective year) and the OK and CANCEL button (starts 3 macros). The graph reflects the values of dependent and independent series. The three necessary values of the independent series are calculated by the spreadsheet functions 'solve for ...' to solve the equation for the assigned values. This done by the macros {ME1_S1},{ME1_S2} and {ME1_S3}. They are executed subsequently leaving the

dialog box with OK and CANCEL. The macro source is in spreadsheet C of the notebook.

The third button: Simply calls the macro, e.g. {GraphView *BA_PIB01*}. The part in italics indicates the graph to show. This is the graph to enter the scenario from, it has a button (<Altere Valores>) which executes the macro {PIB01_M}. This macro branches to the scenario in Spreadsheet B. This macro also copies the values from spreadsheet A used for the scenario graph to spreadsheet C. This prevents the values in the main spreadsheet to be changed. The values entering the scenario graph are all from the spreadsheet C. The data in spreadsheet A remain unaltered.

The fourth button: Calls the macro {AJUDAG3}. This macro branches to the help screen (*UF_ajcenário*). The help screen returns to the scenario with help of the macro {VOLTAG3}. All macros are in spreadsheet C.

The fourth button: Calls the macro {IMPRIM} to print the scenario graph with eventual modifications. The printer used is the standard printer installed under WINDOWS.

For further explications about applications written for QUATTRO PRO see: *QUATTRO PRO, Construindo Aplicativos com Planilhas*. Some helpful examples can be found in the samples subdirectory of QUATTRO PRO Vers 1.0.

Observation: Unfortunately the macros written in QUATTRO PRO V1.0 in Portuguese do not execute in the English Version 5.0 of the same program. Some of the macro statements are language dependent(e.g. {SolveFor.Valor_Destino....}). Whereas 'English' macro statements execute in the Portuguese version, 'Portuguese' macro statements do not in the English version.

8 .Creating the interactive information system

So far as described the system consist of 12 spreadsheets with graphs showing various time series and some scenarios (5 for the time being (PIB, investments, employees, revenues and gains (or value added)) but similar ones can simply be copied and edited with modifies legends and data.

The system TURINFO has to be installed after QUATTRO PRO has been installed previously. The default directory of QUATTRO PRO is C:\QPW. TURINFO for each state should be installed under a unique icon, e.g. the line graph icon of QUATTRO PRO.

The command line of the program item properties has to read like this for the case of TURINFO Bahia:

Q:\QPW\QPW.EXE C:\DBASES\BA_BASE\BA_ENT01.WB1

By executing this line, QUATTRO PRO is initialized and the indicated spreadsheet BA_ENT01.WB1 is loaded.

This spreadsheet contains a macro called {STARTUP}, which has to be executed once the spreadsheet has been loaded. This macro does the following:

- It sets the default properties of QUATTRO PRO, e.g. sets properties necessary for macro execution,
- sets the default subdirectory so that only spreadsheets of this subdirectory are loaded (e.g. DBASES\BA_BASE\..)
- It calls the macro library MACLIB04 which contains all necessary macros
- It executes the slide show to start TURINFO
- It displays the main menu

This macro has once to be executed manually, if QUATTRO PRO has been installed for the first time, e.g. <ALT>+<F2> <STARTUP>. All the times TURINFO is loaded next, it executes automatically, because the default properties have been changed. Since each TURINFO (of a different state) has a different command line (Q:\QPW\QPW.EXE C:\DBASES\UF_BASE\UF_ENT01.WB1, a different macro {STARTUP} is executed by the file UF_ENT01.WB1.

The entry spreadsheet UF_ENT01.WB1 contains one more macro {FIM} which abandons the system WITHOUT SAVING ANY CHANGES. So if you want to save changes do not use this macro to leave QUATTRO PRO.

Apart from this UF_ENT01.WB1 contains only several menu screens showing graph buttons which allow to branch to another screen or select a graph. The whole system is set up as an interactive one, not having to type anything using it. The interaction

between the graphics is done using graph buttons and eventually macros. Graph buttons are graphical element (of a graph) which either call another graph or execute a macro when being pointed at and clicked with a mouse.

All the macros (apart from the ones contained in the scenarios and {FIM} in *UF_ENT01.WB1*) are stored in *MACLIB04.WB1*, which has the property being a macro library according to the standards of QUATTRO PRO.

This file *MACLIB04.WB1* contains some simple macros:

- {SELTAB01} is executed choosing <TABLES>. This opens a dialog box showing all spreadsheet names with states abbreviations and extension *.WB1. This allows entering the spreadsheets to perform manual changes. This option should actually be excluded if users would use the system self contained, inexperienced users might corrupt the system easily.
- {INICIO} closes all open files except *UF_ENT01.WB1* and *MACLIB04.WB1* and returns to the main menu this macro is executed when <MENU> is chosen from a graph of any other spreadsheet than *UF_ENT01.WB1*.
- {INICIO1} returns to the main menu without closing any file, this macro is executed when <MENU> is chosen from a graph of *UF_ENT01.WB1*.

Apart from this there are mainly four different types of macros:

- Initializing a series (e.g. {G_UF_PIB01})
- Initializing a series and call first graph (e.g. {GSUF_PIB01})
- Calling a subsequent screen (e.g. {UF_PIB01_P})
- The concatenated series/graph names indicating which graph has to be executed depending on series and graph (e.g. {SQCOM01_DFPIBC1})

Calling a subsequent screen with <Próximo> executes a macro depending on the series of graphs the macro has been called from, so two names can be considered as parameters (the series name and the graph name). Since in the first version the system was developed in, no true parameter passing was available, this has been simulated by replacing lines in macros and then execute these macros. Graphics series and the corresponding graphs are documented in the document *TURINFO: Introdução ao sistema. Annex L.*)

The macros which execute and concatenate the name of series and graph are

{SPROXIMO} and {SPROXIMOS}. Two macro-names appear to allow for longer graph names to be concatenated, {SPROXIMOS} used for these longer names, e.g. *UF_GHTIN2*³, this also could be consolidated in the future, and use one macro name for this operation only, because {SPROXIMOS} had to be introduced at a later stage of developing TURINFO.

Calling a subsequent graph means to store a graph name in Cell CQ17 of the spreadsheet MACLIB04.WB1.

Initializing a series means to store a series name in Cell CQ8 of the spreadsheet MACLIB04.WB1. {SPROXIMO} and {SPROXIMOS} concatenate (join) these two names which is the name of the macro to be executed next.

An example: The next screen of *UF_PIBC1* is called by macro {*UF_PIBC1_P*}. This macro stores the graph name in Cell CQ17. Relative to the series which has been initialized before by the system, e.g. *GSUF_PIB01* or *GSUF_COM01* the series names *SQ_PIB01* or *SQ_COM01* are stored at Cell CQ8. The macro {SPROXIMO} concatenates these names and executes either macro {*SQ_PIB01_DF_PIBC1*} or {*SQ_COM01_DF_PIBC1*}. This is the fourth group of macros in the library: the concatenated series/graph names indicating which graph has to be executed depending on series **and** graph.

Observation: Modifying MACLIB04.WB1 for a different state and maintaining the structure of the system means to change the file names only, all the graph and series names should remain unchanged, e.g. change all *RJ_* to *AM_* if you want to change the macro library from a copy of the 'Rio de Janeiro'-MACLIB04.WB1 to the 'Amazonas'-MACLIB04.WB1.

9 .Documenting the model and its results

There are mainly three documents to document TURINFO and the results.

- *TURINFO: Introdução ao sistema*
- *TURINFO: Dados Técnicos relativo á*
- *Developing the TURINFO Model, an easy cook-book approach*

³ The macro names in MACLIB04.WB1 contain the letters DF in the name instead of UF. This has not been changed yet and could be consolidated in the future. These correction have not been given high priority because the macro names never appear to the user. The graphs and macro names should not be altered from one system to the other, however UF would be more appropriate than DF. Since the macro names appear in many graphs, as macros being called through graph buttons, to change the macro names is not a mere search and replace in MABLIB04.WB1. Changing the macro names would cause unnecessary maintenance efforts

The first document is self-contained and explains the essentials of TURINFO for the operator and the occasional user. It requires little knowledge as a precondition if used in a workshop as it is intended to. The document is in WordPerfect 6.0 format and is stored in subdirectory C:\WPDOCM in three files MAN*.TXT. Three files were used because of the risk to corrupt files and included graphs during file transfers. All the files have back ups and can easily be edited. The primary editing with **WordPerfect 6.0** for WINDOWS proved to be very difficult. Neither is WordPerfect as reliable as its versions predecessors nor is WINDOWS an reliable working platform for some programs. It is safer to use the DOS version of WordPerfect 6.0, despite of greater comfort of the WINDOWS-Version. Unfortunately some of the table features proved to be incompatible between the DOS and WINDOWS version, so both version were used. Especially difficult it proved to be the import of *.PCX files as graphics. Also the printers used for printout had memory problems and the printing problems can be due to this. The new HP 560C printer works well and reliable, better than the available LASERJET laser printers.

The second document is edited for each state. Several macros have been written to shorten the repetitive task of editing the same type of document. This document consists of four section:

- Tables and explication of variable names
- Results and interpretation of Factorial analysis
- Equations
- Line graphs in the system

For the first section: Use WordPerfect macros “**t1....t25**” and “**tr**” alternating to import part of the TABLE1 in WP format. It is important that only parts of the spreadsheet are imported at one time (usually 7 columns). The macro **tr** which has to be executed immediately after **t1** etc. is removing all the lines from the imported table. Depending of the size of the table all or only some of the macros **t1 .. t25** have to be executed. The file name of the imported table is DF_TAB1.WK1. If using a different file name simply rename the new file (or change the macro **t1**). In the same way import TABLE2 and TABLE3 into WordPerfect. Because these tables are smaller only some of the macros have to be executed, e.g. **t1 to t4** for TABLE2 and **t1 and t2** for TABLE3. Since the end of the tables are identical they have been omitted for TABE2 and TABLE3.

For the second section: The resulting output from factorial analysis have to be cleaned of the unnecessary lines and have to be imported directly into WordPerfect

as an ASCII file . **For the third section:** Saving the output files of regression analysis of the statistics program as ASCII files (the equations), these can be imported directly into WordPerfect. It is more convenient to make spreadsheets of the equations belonging to one subject, e.g. PIB or investment and import the spreadsheet into WordPerfect. The table feature maintains the output format even using then better looking proportional spaced font. If using the ASCII files a non-proportional spaced font (like COURIER) has to be used to maintain the table structure of the equation output.

There have been some requests from TURINFO users to include Regression fit (R^2) and significance levels as well as Durbin-Watson values in the output. This would unnecessarily inflate the document. Since in an iterative process the regression coefficients are calculated, output can be very large for each of these calculations. With the name of the included variables at hand a user experienced in using statistical software can repeat the calculation with ease. Use WordPerfect macro **tb1** to erase all lines from the imported tables.

For the fourth section: The graphs are results of exports from QUATTRO PRO. Use the spreadsheet EXP03.WB1 which contains macro lines. Run the macro from the first line of spreadsheet EXP03.WB1 of the *UF_DOCU* subdirectory to produce all of the output graphics in *.PCX format (if done manually this can be quite time-consuming). The enter WordPerfect (for DOS) and execute macros **rep1UF** . to . **rep9UF**. This will produce 9 output files *UF_PIB.GRA*, *UF_NESAQ.GRA*, *UF_AQUPE.GRA*, *UF_PERSA.GRA*, *UF_INPDE.GRA*, *UF_DESRE.GRA*, *UF_RECVA.GRA*, *UF_VALGA.GRA* and *UF_TURIS.GRA*.

An existing spreadsheet C:\DBASES\PR_BASE\PR_DOCU\DF_NESAQ.GRA is used in all the macros to substitute the graphics by the actual ones, so this spreadsheet should not be deleted. The files with extension *.GRA can be printed normally they have preset line numbering from Page 50 to Page 84.

10.Making TURINFO work

The training aspect is quite important making TURINFO be used by a large number of people. Obviously a person with statistical and economical background has less difficulty to understand and interpret equations and dependencies. The main target group are staff members of tourist organizations and their clients. The clients in this

case are not necessarily tourists but entrepreneurs, planners and several professional organizations, like ABAV, SEBRAI etc. To make those people interested in the first place is to make them understand the system. It showed that the understanding is not at all that straightforward for a person not trained in statistical thinking. It is easy however to handle and use the system, even for a complete layman. To bridge the gap between the person using the system without understanding much of the meaning of the graphs and the content of the system, which can be quite rich in information, it needs interpreters.

These interpreters were thought to be staff members of the tourist offices and EMBRATUR.

The other important aspect of making TURINFO work in the future is the availability of reliable data and the access to it in the future. One essential shortcoming is the difficulty of access to IBGE data. The data of the Statistical Institute seem to be of good quality and rather reliable. One problem is the missing census of 1990, the other is the problem of dissemination of data. It seems to be that technical problems make it impossible for IBGE to answer many of the requests to furnish data. The necessity of dissemination data on diskettes, even better would be a mailbox-service, is ever increasing. Unfortunately IBGE responds only reluctantly to these requests, maybe more because of financial reasons than because of anything else. Data base projects of EMBRATUR (SINTUR in Brasília and a DEMAND data base being developed in Rio) might respond to the data problem on the tourist demand side, but the main burden of the data problem is on the shoulders of IBGE. Moreover some essential enquiries seem to be missing for implementing a system like TURINFO with more reliability. No household survey as been carried out on a national level for more than 10 years (this has of cause much more dramatic aspects than the lack of tourism data). No nationwide survey has been carried out on national tourist demand since 1987, which is probably much more important for the national tourism industry than staring at the foreign tourist⁴, with Brazil having a GDP of almost the double than the rest of South America (Source: World Bank 1990) and a population of about the same as the rest of South America (Source: STBA, Stat.Jahrbuch 1992 f.d. Ausland). Having better economical data the Satellite Account approach might be useful too, but it seems there will be limits as to apply National Accounts concepts to regional structures as the Brazilian States. It seems rather necessary to combine the two approaches, as to find out about the states

⁴A 1986/87 survey estimated the number of Brazilian tourists to be 42 millions in that year against 1.5 millions of foreign tourists.

economies in general and about Tourism and its participation in the states economies in detail.

Annex

Location of files on the project computer and the backup tape. The backup tape was written with EZTAPE™ Software. It is strongly recommended that regular backup copies are maintained of the TURINFO files.

TURINFO model files: C:\DBASES\UF_BASE (see page 9)

Documentation of each model: C:\DBASES\UF_BASE\UF_DOCU, including *TURINFO:Dados técnicos relativo...*, e.g. RJ_MT_P1.DOC (includes part 2, Factorial analysis) and RJ_MT_P3.DOC for the first three parts of this document. The graphs are stored in compacted files, e.g. PR_PJ.ARJ under subdirectory C:\T1 and have to be extracted before they can be altered or printed in **WordPerfect 6.0**. Title pages are stored under the same directory, e.g. PAGTIT2I.WPW and can be printed under **WordPerfect 6.0**.

General documentation: All general documents are stored under C:\WPDOCS like MAN-20.TXT, MAN21-.TXT and MANANX1.TXT, three files to store the document *TURINFO:Introdução ao sistema*. Also the title pages are stored there (PAGTIT1F.WPW) as well as other documents (Cook book: DEVMOD01.DOC etc.). Title pages have the extension CDR if they are CorelDraw files.

WP Macros: are stored under C:\WP60\MACROS