

Community Innovation Survey (CIS) microdata

Note to the CIS researcher visiting Eurostat's SAFE Centre

15.7.2013 Eurostat – Unit G6 Innovation and information society

I Introduction

An authorised researcher has the right to use via the Safe Centre at Eurostat the Community Innovation Survey microdata for an agreed research project. This possibility depends on overall microdata availability at Eurostat (CIS microdata provisions are voluntary), Member States' willingness to allow the CIS microdata to be offered for the research use and the permission for using the data for the particular research project.

Before entering to the SAFE Centre the researcher has received and read the Eurostat Manual on the protection of confidential data; has signed the Commitment on statistical confidentiality; has received service card and personal access card for the SAFE Centre; has been furnished by the user ID and a password and has been personally briefed by Eurostat's legal service and the Unit in charge of the CIS data, Unit Innovation and information society.

II Eurostat SAFE Centre

Visiting times are the standard office times from 7H00 to 20H00. Core office hours are 9H15 to 16H45. During the core hours the assistance on PC, Eurostat's software, CIS data, etc. is guaranteed. The beginning of the first visiting day of a researcher needs also to be fitted in the core hours. Software procedures can be left running overnight although there are no guarantees on the problem free execution due to potential power cuts, network updates and so on.

Researcher works within a given directory of a stand-alone PC. The source data is stored under the directory which is dedicated only for the researcher and all the programs and the output will be stored in that directory. Duplicating of the source data sets or subsets of it shall be avoided for not to extend unnecessary the volume of the stored data.

Should there be another researcher coming for the stand alone PC before the work of the previous researcher is finished, all the work files (data, programs, interim results,..) will be moved away from the PC as it will be fully cleaned (reset) for the new user. When the first researcher continues the work the software needs to be re-installed (by the researcher) and the work files needs to be re-loaded (by Eurostat). When working consecutive days or without interruption by another researcher (within short period) this will not be necessary.

The available software is Microsoft office 2003, SAS 9.2 with SAS Enterprise Guide SAS 4.3 and STATA 12 (basic version).

In the SAFE Centre it is **NOT** possible

- to print documents;
- to copy data to CD-ROM, DVD, Zip drives or USB keys;
- to copy data to the local hard disk (in secure server environment);
- to connect recording devices to the serial, parallel and USB ports;
- to connect a laptop to the network;
- to use email;

- to make Internet connections.

III Researcher's own data and own software

Should the researcher wish to use own data set(s), this will not happen without Eurostat's explicit prior agreement. If the use of the researcher's own data set is indispensable for the project it is advised to contact Eurostat on that before delivering the application of the research project. Own data may concern for example background variables such as national totals or particular variables for separate economic activities.

Any data or software of the researcher needs to be stored in CD-ROM or in USB-key (used ports only possible to read) to be able to be loaded into the stand alone PC. Researcher her-/himself is in charge of the functionality of own software.

IV CIS data sets

The countries and years covered are indicated in the research contract.

The data sets are provided for researcher's use either in CSV or STATA format. Particular notes on the microdata set will be provided. CIS 2010 data structure and data coding is annexed.

V Output validation

The researcher should ensure that any results of the research published or otherwise disseminated do not contain information which may permit the identification of individual records of the data.

Any deliberate attempt to compromise the confidentiality of persons or organisations to which confidential data for scientific purpose related may result in prosecution in accordance with applicable law.

All the results to be transmitted away from the SAFE Centre are nevertheless checked by Eurostat to avoid any disclosure of confidential data.

The non-identification covers both primary and secondary confidentiality.

Primary confidentiality concerns tabular cell data, whose dissemination would permit attribute disclosure. The two main reasons for declaring data to be primary confidential are:

- Too few enterprises in a cell;
- Dominance of one or two enterprises in a cell.

Secondary confidentiality concerns data which is not primary disclosive, but whose dissemination, when combined with other data permits the identification of an enterprise or the disclosure of an attribute of the enterprise.

Further to the above, any statistics (tables, graphs, textual references) on any kind of sub-population (cell) shall not be published:

- (1) if they consist of less than 10 enterprises;
- (2) where one enterprise represents more than 70% of the total sub-population expenditures, employment or turnover;

(3) where two enterprises represent more than 85% of the total sub-population expenditures, employment or turnover.

In addition, where there are primary confidential cells, the secondary confidentiality treatment is necessary to make sure that these primary confidentiality cells cannot be estimated with the help of the other non-confidential cells.

Even if the confidentiality has been defined to be in the researcher's (or researcher's background institution's) responsibility, Eurostat validates all output the researcher wishes to export from the SAFE Centre. Researcher shall be able to explain the processes and show that the output is non-disclosive. Data which has been validated is safe to use further outside the SAFE Centre or to publish as such.

In the following there are some illustrative examples on few situations as regards the confidentiality. Examples are drawn from another enterprise based dataset of Eurostat, namely Structure of Earnings Survey data (SES). The examples and remarks below are given only as an indication and guideline for researcher. They do not cover all the possible situations and possibilities. General non-disclosure principle and the rules specified above shall be respected in all circumstances.

Example of primary confidentiality

Primary confidentiality means that any cell of the output to be exported from the SAFE Centre needs to fulfil directly the conditions above (1 to 3).

Illustration (i)

Region	Economic activity	Occupation	Median earnings	Number of local units	Number of employees
AA1	61	21	25.2	20	120
AA1	61	22	30.5	4	25
AA1	61	23	22.2	18	55
AA1	61	24	24.4	16	210
AA1	61	25	19.1	31	482
Total AA1	61	21-26	23.1	89	892

In the Illustration (i), the occupation 22 does not fulfil the condition that the published cell shall have 10 or more units in it. This output proposal would be rejected as the number of local units in occupation 22 is only 4.

Example secondary confidentiality

Hiding the occupation 22 in the Illustration (i) would create a problem of secondary confidentiality: a reader would be able to calculate the number of the local units in the hidden cell using the total AA1 and non-hidden information. Also the hidden sensitive information, median earnings, could become easily estimated for the occupation 22, at least its range. This output proposal would be rejected even with hiding the information in the line of the occupation 22.

Illustration (i) is just one relatively simple example of the secondary confidentiality. While protecting and validating the secondary confidentiality, the data in different *independent* tables and different forms of presentations (graphs) and *classification levels* and systems shall be taken into account (some aspects to list).

Example of dominance

Alike with the rule of having at least 10 units in the published cell, the output must fulfil the dominance rules 2 and 3.

As the SES data set does not have amongst mandatory variables any precise measure on the size of the enterprise or the local unit, dominance rules are linked to the number of employees in each unit and the gross earnings they represent.

Illustration (ii)

Local unit/ Enterprise	Average gross earnings in €	Number of employees	Total gross earnings in €
1	2 333	5	11 665
2	3 535	603	2 131 605
3	2 802	10	28 020
4	2 956	12	35 472
5	1 999	6	11 994
6	2 716	10	27 160
7	2 350	9	21 150
8	2 752	20	55 040
9	2 232	6	13 392
10	1 998	10	19 980
Total 1-10	3 409	691	2 355 478

The whole table in the Illustration (ii) represents one cell. There only the total Average gross earnings € 3 409 was aimed to be published. However as the second local unit / enterprise represents 87% of the employees and 90% of the gross earnings, the total Average gross earnings € 3 409 of this cell cannot be published. The dominance rule with two largest units of the cell works similarly. Further, the secondary confidentiality needs to be respected also in the dominance rule context.

Regressions and other than tabular forms of output

Because of the secondary confidentiality, the tabular data output may become very complicated to validate. This may also be the situation when moving further from the standard tables. Linear and non-linear estimation, simulation, modelling, different types of developed analysis, particular indices and all (other) kind of econometric methods and their output may require a lot of specific knowledge to be able to validate the disclosiveness of the output.

In general, regression results are non-disclosive at an exact level (some inferences may be drawn within a margin of error in particular cases). Moreover, this small risk can be reduced further in ways which do not significantly reduce the usefulness of the results. The simplest way is non-reporting of incidental parameters, such estimated constants or the coefficients on irrelevant dummy variables. In general a regression with $(N-K) \rightarrow \infty$ which does not report all significant parameters is non-disclosive for all practical purposes.

For other analytical results, the disclosive nature depends on the manipulations carried out. The assumption is that results are disclosive unless proved otherwise, and therefore it is in researcher's interests to show that the results are non-disclosive.

Graphs are also treated as tables which just present the information in a different form. *Quantiles* are also tables. *Maximum* and *minimum* values are also tables with normally only one enterprise in a cell i.e. they are confidential and cannot therefore be released.

Detecting and protecting secondary confidentiality also for the other than tabular forms of output shall be ensured.

Rejection of the output

The proposed output will be automatically rejected if the rules (1) to (3) above are not respected. The output may also be rejected if it is not fully understood or the output is very long. In these cases Eurostat cannot be sure whether the confidentiality rules are fully respected and cannot therefore validate the data for certainty. Non-documented or unexplained output will not be approved (presentations of tables or other results alone).

Eurostat does not necessary make proposals how to modify the output to get it accepted but just indicates the reason why the output has been rejected. If the output is rejected researcher needs to re-work the output for having it re-validated.

Notice that the burden of proving that results are safe is onto the researcher. Further, it is in researcher's own interest to show the output which can be validated within reasonable time and without potential costly re-visits at SAFE Centre.

Preparing the output for validation

The proposed output shall include all the information which is needed for the output validation even if this (extra) information will not be published / used further (frequencies etc.). For example, the two last columns in the Illustration (i) are necessary for validating the data even they would not be published in the final report the researcher will prepare. Notice also that the two last columns would NOT be enough for Eurostat to validate the dominance rules (2-3) in the Illustration (i). On the other hand Illustration (ii) includes the necessary information for validating the dominance rules. It is up to the researcher to decide the type of presentation and measures in showing that all the rules have been respected.

Together with the results to be validated, all the programs to derive this output will have to be presented as all the results must be reproducible.

It should be noticed that the output validation concerns only the disclosiveness of the data. Output validation is not a quality check. Appropriateness of the assumptions or underlying theory or analysis will not be assessed nor the conclusions drawn. All this remains researcher's responsibility.

VI Output validation form and time

The output wished to be exported from the SAFE Centre needs to be saved (together with the related programs) under given directory. The table data shall be saved together with all the other data necessary for the validation with the adequate headings, titles and other metadata. New derived variables should be documented and meaningful variable names have to be used.

At the end of the research work, researcher informs Eurostat Unit G6 contact on the wish to export the data from the SAFE Centre after which the validation process begins. It is recommended that the researcher in person shows the output to Eurostat before leaving the SAFE Centre and explain the main characteristics for facilitating the validation work.

Eurostat reserves to itself the right to define the time needed for validating the output data. All technical and organisational measures will be taken by Eurostat to ensure an efficient checking without undue delays. Eurostat will make every effort to ensure that this delay does not exceed two weeks.

The validated output shall be sent to the researcher by email.

Read more on output checking from [http://www.cros-portal.eu/content/work-packages-2:Deliverables_Task 3'](http://www.cros-portal.eu/content/work-packages-2:Deliverables_Task%203)

ANNEX

CIS 2010 DATA STRUCTURE AND CODING – Forewords

1. The order of the variables and the codes used correspond to the order and the coding of the Harmonised Survey Questionnaire of the Community Innovation Survey 2010 (FINAL VERSION July 9, 2010).
2. All monetary variables are in Euro. There are no separator for thousands, millions etc. For non-Euro area currencies, the exchange rates used for 2008 or 2010 are annual averages.
3. All percentages have been expressed as proportions and rounded to the nearest figures (e.g. 25.5% have been expressed as 0.26). Therefore 1.00 would be equivalent to 100%.
4. A point has been used for NACE (XX•XX) i.e. 10.81.
5. The weighting factors WEIGHT and WEIGHTNR are rounded up to three decimal points (XXX•XXX) i.e. 7.236. The variable WEIGHT is the original weighting or grossing factor. If non-response survey has not been conducted (i.e. the response rate was sufficiently high) there will be no correction for non-response bias. If there was a correction to the weights for bias detected in the non-response survey, this final weight has been recorded as WEIGHTNR. More weighting factors may have been used for some countries (explanatory notes will be offered).
6. Each CIS data set represents independent enterprise sample. Independently of the fact that several enterprises are sampled consecutively from one CIS wave to another (particularly the largest enterprises) the CIS data sets do not represent enterprise panel. Enterprise identifiers do not allow linking of the enterprises over CIS waves.

CIS 2010 DATA STRUCTURE AND CODING – Variables

Variable	Length	Question	Code
General information about the enterprise			
ID	6	Non-informative record identifier	XXXXXX
NUTS	4	Country code	NUTS level0 (country)
NACE	5	Main activity	NACE Rev. 2 XX.XX i.e. 29.31
GP	1	Enterprise part of a group	0:No 1:Yes • : missing
HO	2	Country of head office	ISO 3166 alphanumeric 2 digit code • : missing
MARLOC	1	Local/regional market (within country)	0: No 1: Yes • : missing.
MARNAT	1	National market (other regions of country)	0: No 1: Yes • : missing.
MAREUR	1	Other EU/EFTA/CC market	0: No 1: Yes • : missing.
MAROTH	1	All other countries	0: No 1: Yes • : missing.
LARMAR	1	Largest market in terms of turnover between 2008 and 2010	A. Local / regional within [your country] B. National (other regions of [your country]) C. Other European Union (EU), EFTA, or EU candidate countries D. All other countries
Product innovation			
INPDGD	1	Introduced onto the market a new or significantly improved good	0:No 1:Yes • : missing
INPDSV	1	Introduced onto the market a new or significantly improved service	0:No 1:Yes • : missing
Development of goods innovation			
INPDTG1	1	Your enterprise by itself	0:No 1:Yes • : missing
INPDTG2	1	Your enterprise together with other enterprises or institutions	0:No 1:Yes • : missing
INPDTG3	1	Your enterprise by adapting or modifying goods or services originally developed by other enterprises or institutions	0:No 1:Yes • : missing
INPDTG4	1	Other enterprises or institutions	0:No 1:Yes • : missing
Development of service innovation			
INPDTS1	1	Your enterprise by itself	0:No 1:Yes • : missing
INPDTS2	1	Your enterprise together with other enterprises or institutions	0:No 1:Yes • : missing
INPDTS3	1	Your enterprise by adapting or modifying goods or services originally developed by other enterprises or institutions	0:No 1:Yes • : missing
INPDTS4	1	Other enterprises or institutions	0:No 1:Yes • : missing
Novelty level of product innovation			
NEWMKT	1	Did the enterprise introduce a product new to the market	0:No 1:Yes • : missing
NEWFRM	1	Did the enterprise introduce a product new to the firm	0:No 1:Yes • : missing
TURNMAR	4	% of turnover in new or improved products introduced during 2008-2010 that were new to the market	0.00 to 1.00 • : missing
TURNIN	4	% of turnover in unchanged or marginally modified products during 2008-2010 that were new to the firm	0.00 to 1.00 • : missing
TURNUNG	4	% of turnover in unchanged or marginally modified products during	0.00 to 1.00 • : missing

Variable	Length	Question	Code
2008-2010			
Geographical level of novelty for product innovation			
INPDFC	1	A first in [your country]	0:No 1:Yes 2: Don't know • : missing
INPDFE	1	A first in Europe	0:No 1:Yes 2: Don't know • : missing
INPDFW	1	A world first	0:No 1:Yes 2: Don't know • : missing
Process innovation			
INPSPD	1	Introduced onto the market a new or significantly improved method of production	0:No 1:Yes • : missing
INPSLG	1	Introduced onto the market a new or significantly improved logistic, delivery or distribution system	0:No 1:Yes • : missing
INPSSU	1	Introduced onto the market a new or significantly improved supporting activities	0:No 1:Yes • : missing
Development of process innovation			
INPSDV1	1	Your enterprise by itself	0:No 1:Yes • : missing
INPSDV2	1	Your enterprise together with other enterprises or institutions	0:No 1:Yes • : missing
INPSDV3	1	Your enterprise by adapting or modifying processes originally developed by other enterprises or institutions	0:No 1:Yes • : missing
INPSDV4	1	Other enterprises or institutions	0:No 1:Yes • : missing
Novelty level of process innovation			
INPSNM	1	Were any of your process innovations introduced between 2008 and 2010 new to your market	0:No 1:Yes 2 : Do not know • : missing
Not yet completed or abandoned activities (for product and/or process innovation)			
INABA	1	Abandoned or suspended before completion	0:No 1:Yes • : missing
INONG	1	Still ongoing at the end of the 2010	0:No 1:Yes • : missing
Innovation activities and expenditures (for product and/or process innovation)			
RRDIN	1	Engagement in intramural R&D	0: No 1: Yes • : missing.
RDENG	1	Type of engagement in R&D	1: continuously 2: occasionally • : missing
RRDEX	1	Engagement in extramural R&D	0: No 1: Yes • : missing.
RMAC	1	Engagement in acquisition of machinery	0: No 1: Yes • : missing.
ROEK	1	Engagement in acquisition of external knowledge	0: No 1: Yes • : missing.
RTR	1	Engagement in training for innovative activities	0: No 1: Yes • : missing.
RMAR	1	Engagement in market introduction of innovation	0: No 1: Yes • : missing.
RDSG	1	Design	0: No 1: Yes • : missing.
RPRE	1	Engagement in other preparation	0: No 1: Yes • : missing.
RRDINX	8	Expenditure in intramural R&D	0 to 99999999 • : missing
RRDEXX	8	Purchase of extramural R&D	0 to 99999999 • : missing
RMACX	8	Expenditure in acquisition of machinery	0 to 99999999 • : missing
ROEKX	8	Expenditure in acquisition of external knowledge	0 to 99999999 • : missing

Variable	Length	Question	Code
RTOT	8	Total of these four innovation expenditure categories	0 to 99999999 • : missing
Public funding of innovation (for product and/or process innovation)			
FUNLOC	1	Public funding from local or regional authorities	0:No 1:Yes • : missing
FUNGMT	1	Public funding from central government	0:No 1:Yes • : missing
FUNEU	1	Public funding from the EU	0:No 1:Yes • : missing
FUNRTD	1	Funding from EU's 7th Framework Programme for RTD	0:No 1:Yes • : missing
Sources of information for innovation (for product and/or process innovation)			
SENTG	1	Sources from within the enterprise or enterprise group	0: Not used. 1: Low 2: Med 3: High • : missing.
SSUP	1	Sources from Suppliers of equipment, materials, etc.	0: Not used. 1: Low 2: Med 3: High • : missing
SCLI	1	Sources from Clients or customers	0: Not used. 1: Low 2: Med 3: High • : missing
SCOM	1	Sources from Competitors and other enterprises of same industry	0: Not used. 1: Low 2: Med 3: High • : missing
SINS	1	Sources from consultants, commercial labs or private R&D institutes	0: Not used. 1: Low 2: Med 3: High • : missing
SUNI	1	Sources from Universities or other higher education institutes	0: Not used. 1: Low 2: Med 3: High • : missing
SGMT	1	Sources from Government or public research institutes	0: Not used. 1: Low 2: Med 3: High • : missing
SCON	1	Sources from professional conferences, trade fairs, meetings	0: Not used. 1: Low 2: Med 3: High • : missing
SJOU	1	Sources from Scientific journals, trade/scientific publications	0: Not used. 1: Low 2: Med 3: High • : missing
SPRO	1	Sources from Professional and industry associations	0: Not used. 1: Low 2: Med 3: High • : missing
Innovation cooperation (for product and/or process innovation)			
CO	1	Cooperation arrangements on innovation activities	0:No 1:Yes • : missing
CO11	1	Other enterprises within enterprise group : National	0:No 1:Yes • : missing
CO12	1	Other enterprises within enterprise group : EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO13	1	Other enterprises within enterprise group : US	0:No 1:Yes • : missing
CO14	1	Other enterprises within enterprise group : China or India	0:No 1:Yes • : missing
CO15	1	Other enterprises within enterprise group : Other countries	0:No 1:Yes • : missing
CO21	1	Suppliers of equipment, etc. : National	0:No 1:Yes • : missing
CO22	1	Suppliers of equipment, etc. : EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO23	1	Suppliers of equipment, etc. : US	0:No 1:Yes • : missing
CO24	1	Suppliers of equipment, etc. : China or India	0:No 1:Yes • : missing
CO25	1	Suppliers of equipment, etc. : Other countries	0:No 1:Yes • : missing
CO31	1	Clients or customers : National	0:No 1:Yes • : missing
CO32	1	Clients or customers : EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO33	1	Clients or customers : US	0:No 1:Yes • : missing
CO34	1	Clients or customers : China or India	0:No 1:Yes • : missing
CO35	1	Clients or customers : Other countries	0:No 1:Yes • : missing
CO41	1	Competitors or other firms ... : National	0:No 1:Yes • : missing
CO42	1	Competitors or other firms ... : EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO43	1	Competitors or other firms ... : US	0:No 1:Yes • : missing
CO44	1	Competitors or other firms ... : China or India	0:No 1:Yes • : missing

Variable	Length	Question	Code
CO45	1	Competitors or other firms ... : Other countries	0:No 1:Yes • : missing
CO51	1	Consultants, commercial labs, private R&D institutes : National	0:No 1:Yes • : missing
CO52	1	Consultants, commercial labs, private R&D institutes : EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO53	1	Consultants, commercial labs, private R&D institutes : US	0:No 1:Yes • : missing
CO54	1	Consultants, commercial labs, private R&D institutes : China or India	0:No 1:Yes • : missing
CO55	1	Consultants, commercial labs, private R&D institutes : Other countries	0:No 1:Yes • : missing
CO61	1	Universities or other ... : National	0:No 1:Yes • : missing
CO62	1	Universities or other ... : EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO63	1	Universities or other ... : US	0:No 1:Yes • : missing
CO64	1	Universities or other ... : China or India	0:No 1:Yes • : missing
CO65	1	Universities or other ... : Other countries	0:No 1:Yes • : missing
CO71	1	Government or public research institutes: National	0:No 1:Yes • : missing
CO72	1	Government or public research institutes: EU/EFTA/EU-CC	0:No 1:Yes • : missing
CO73	1	Government or public research institutes: US	0:No 1:Yes • : missing
CO74	1	Government or public research institutes: China or India	0:No 1:Yes • : missing
CO75	1	Government or public research institutes: Other countries	0:No 1:Yes • : missing
PMOS	1	Most important co-operation partner	A: Other enterprises within enterprise group B: Suppliers of equipment, etc C: Clients or customers D: Competitors or other firms E: Consultants, commercial labs, private R&D institutes F: Universities or other ... G: Government or public research institutes
Innovation objectives (for product and/or process innovation)			
ORANGE	1	Increased range of goods or services	0: Not relevant 1:Low 2:Med 3:High • : missing
OREPL	1	Replace outdated products or processes	0: Not relevant 1:Low 2:Med 3:High • : missing
ONMOMS	1	Enter new markets or increase market share	0: Not relevant 1:Low 2:Med 3:High • : missing
OQUA	1	Improve quality of goods or services	0: Not relevant 1:Low 2:Med 3:High • : missing
OFLEX	1	Improve flexibility for producing goods or services	0: Not relevant 1:Low 2:Med 3:High • : missing
OCAP	1	Increase capacity for producing goods or services	0: Not relevant 1:Low 2:Med 3:High • : missing
OLBR	1	Reduce labour costs per unit output	0: Not relevant 1:Low 2:Med 3:High • : missing
ORME	1	Reduce material and energy costs per unit output	0: Not relevant 1:Low 2:Med 3:High • : missing
OREI	1	Reduce environmental impacts	0: Not relevant 1:Low 2:Med 3:High • : missing
OHESY	1	Improve health or safety of your employees	0: Not relevant 1:Low 2:Med 3:High • : missing
Factors hampering innovation			
HFENT	1	Lack of funds within your enterprise or group	0: Not experienced 1:Low 2:Med 3:High • : missing
HFOUT	1	Lack of finance from sources outside your enterprise	0: Not experienced 1:Low 2:Med 3:High • : missing
HCOS	1	Innovation costs too high	0: Not experienced 1:Low 2:Med 3:High • : missing

Variable	Length	Question	Code
HPER	1	Lack of qualified personnel	0: Not experienced 1:Low 2:Med 3:High • : missing
HTEC	1	Lack of information on technology	0: Not experienced 1:Low 2:Med 3:High • : missing
HINF	1	Lack of information on markets	0: Not experienced 1:Low 2:Med 3:High • : missing
HPAR	1	Difficulty in finding cooperation partners for innovation	0: Not experienced 1:Low 2:Med 3:High • : missing
HDOM	1	Market dominated by established enterprises	0: Not experienced 1:Low 2:Med 3:High • : missing
HDEM	1	Uncertain demand for innovative goods or services	0: Not experienced 1:Low 2:Med 3:High • : missing
HPRIOR	1	No need due to prior innovations by your enterprise	0: Not experienced 1:Low 2:Med 3:High • : missing
HMAR	1	No need because of no demand for innovations	0: Not experienced 1:Low 2:Med 3:High • : missing
Organisational information			
ORGBUP	1	New business practices for organising procedures	0:No 1:Yes • : missing
ORGWKP	1	New methods of organising work responsibilities and decision making	0:No 1:Yes • : missing
ORGEXR	1	New methods of organising external relations	0:No 1:Yes • : missing
Objectives of organisational innovation			
ORORED	1	Reduce time to respond to customer or supplier needs	0: Not relevant 1:Low 2:Med 3:High • : missing
OROABL	1	Improve ability to develop new products or processes	0: Not relevant 1:Low 2:Med 3:High • : missing
OROQUA	1	Improve quality of your goods or services	0: Not relevant 1:Low 2:Med 3:High • : missing
ORORCO	1	Reduce costs per unit output	0: Not relevant 1:Low 2:Med 3:High • : missing
OROCIN	1	Improved communication or information sharing	0: Not relevant 1:Low 2:Med 3:High • : missing
Marketing information			
MKTDGP	1	Significant changes to the aesthetic design or packaging	0:No 1:Yes • : missing
MKTPDP	1	New media or techniques for product promotion	0:No 1:Yes • : missing
MKTPDL	1	New methods for product placement or sales channels	0:No 1:Yes • : missing
MKTPRI	1	New methods of pricing goods or services	0:No 1:Yes • : missing
Objectives of marketing innovation			
OMKTS	1	Increase or maintain market share	0: Not relevant 1:Low 2:Med 3:High • : missing
OMKTCG	1	Introduce products to new customer groups	0: Not relevant 1:Low 2:Med 3:High • : missing
OMKTGM	1	Introduce products to new geographic markets	0: Not relevant 1:Low 2:Med 3:High • : missing
Creativity and skills available in the enterprise			
SGALA	1	Graphic arts / layout / advertising	0: Skills not used / not relevant 1: Employed in-house 2: Obtained from external sources 3: Employed in-house AND Obtained from external sources • : missing
SDOS	1	Design of objects or services	0: Skills not used / not relevant 1: Employed in-house 2: Obtained from external sources 3: Employed in-house AND Obtained from external sources • : missing
SMED	1	Multimedia (combining audio, graphics, text, still pictures, animation, video .etc)	0: Skills not used / not relevant 1: Employed in-house 2: Obtained from external sources 3: Employed in-house AND Obtained from external sources • : missing
SWDS	1	Web design	0: Skills not used / not relevant 1: Employed in-house

Variable	Length	Question	Code
SSWD	1	Software development	2: Obtained from external sources 3: Employed in-house AND Obtained from external sources •: missing 0: Skills not used / not relevant 1: Employed in-house
SMKR	1	Market research	2: Obtained from external sources 3: Employed in-house AND Obtained from external sources •: missing 0: Skills not used / not relevant 1: Employed in-house
SENAP	1	Engineering / applied sciences	2: Obtained from external sources 3: Employed in-house AND Obtained from external sources •: missing 0: Skills not used / not relevant 1: Employed in-house
SMSDM	1	Mathematics / statistics / database management	2: Obtained from external sources 3: Employed in-house AND Obtained from external sources •: missing 0: Skills not used / not relevant 1: Employed in-house
Methods to stimulate new ideas or creativity among the staff			
MBRST	1	Brainstorming sessions	0: Method not used 1: Successful 2: Not Successful 3: Don't know if successful •: missing
MMDCF	1	Multidisciplinary or cross-functional work teams	0: Method not used 1: Successful 2: Not Successful 3: Don't know if successful •: missing
MJBRT	1	Job rotation of staff to different departments or other parts of your enterprise group	0: Method not used 1: Successful 2: Not Successful 3: Don't know if successful •: missing
MFIN	1	Financial incentives for employees to develop new ideas	0: Method not used 1: Successful 2: Not Successful 3: Don't know if successful •: missing
MNFIN	1	Non-financial incentives for employees to develop new ideas, such as free time, public recognition, more interesting work, etc	0: Method not used 1: Successful 2: Not Successful 3: Don't know if successful •: missing
MTREM	1	Training employees on how to develop new ideas or creativity	0: Method not used 1: Successful 2: Not Successful 3: Don't know if successful •: missing
Basic information on the enterprise			
TURN08	9	Total turnover in 2008	-999999999 to 999999999 •: missing
TURN10	9	Total turnover in 2010	-999999999 to 999999999 •: missing
EMP08	6	Total number of employees in 2008	0 to 999999 •: missing
EMP10	6	Total number of employees in 2010	0 to 999999 •: missing
Percent of enterprise's employees in 2010 with a university degree			
EMPUD	1	% of employees in 2010 with a university degree	0: 0% 1: 1% to 4% 2: 5% to 9% 3: 10% to 24% 4: 25% to 49% 5: 50% to 74% 6: 75% to 100%
Sampling information			
Weight	8	Original weight	
Weightnr	8	Corrected weight (corrected for non-response bias)	