

**Tertiary  
Education  
Commission**

*Te Amorangi  
Mātauranga Matua*



# Performance-Based Research Fund

2018 Quality Evaluation  
Staff Data File specification

Updated: May 2018

Published by the Tertiary Education Commission

Te Amorangi Mātauranga Matua

National Office

44 The Terrace

PO Box 27048

Wellington, New Zealand

Updated May 2018

### Authors

The Tertiary Education Commission

Every effort is made to provide accurate and factual content. The TEC, however, cannot accept responsibility for any inadvertent errors or omissions that may occur.



This work is licensed under the Creative Commons Attribution 4.0 International licence. You are free to copy, distribute, and adapt the work, as long as you attribute the work to the Tertiary Education Commission and abide by the other licence terms. Please note you may not use any departmental or governmental emblem, logo, or coat of arms in any way that infringes any provision of the Flags, Emblems, and Names Protection Act 1981.

## Changes August 2017

Change	Reference
1. Position Title field moved to position 9 after Family Name	Pages 3, 6
2. Tertiary education organisation names and gender codes updated	Appendix Pages 10,11
3. Corrected Field 'Position Title' to Cardinality of 1	Page 6

## Changes May 2018

Change	Reference
1. Corrected Field 'Position Title' to Cardinality of 1	Page 6

## Introduction

This document describes the fields that make up the Performance-Based Research Fund (PBRF) Staff Data file for the PBRF 2018 Quality Evaluation.

The data submitted by tertiary education organisations (TEOs) must include details of all staff members employed or contracted for services between 15 June 2017 and 14 June 2018:

- › for whom an EP has been submitted for the 2018 Quality Evaluation
- › who are PBRF eligible but left the TEO between 15 June 2017 and 14 June 2018 (potentially transferring staff)
- › who are PBRF eligible and concurrently employed by another TEO at 14 June 2018 that is submitting the EP.

This collection of staffing data is required to support the assessment and audit of the 2018 Quality Evaluation only.<sup>1</sup>

The audience for this document includes:

- › TEO developers involved in the development of systems to supply staff data to the TEC
- › TEO business users who need to understand the data items to export using human resource systems.

While this document has been written with a non-technical audience in mind it does assume a basic understanding of comma-separated value (CSV) format.

This document describes the data elements contained within the staff data collection, their meaning, data types, and basic validation rules.

## Terminology and conventions

This document describes data to be supplied in a CSV format. Some basic terminology and conventions used in this document are:

- › Data Field – An item of data to be supplied is known as a "field", e.g. LastName and FirstName are examples of fields.
- › Comma and Quote delimited – Field values are normally enclosed within an open and close double quote (e.g. "field value") and each field is separated by a comma ("field1","field2", "field3").
- › Using quote marks allows the use of more complex data values where commas are part of the data value and not to be confused with the comma to separate fields (e.g. "field1","field, 2, value","field3"). The value for the second field in the previous example is "field, 2, value" and will be interpreted as a single value. Empty quotes (e.g. "", "field2") will be interpreted as a null or an empty or not supplied value.
- › Data Types – This document lists the type or format of data required for a given data field (e.g. string, integer, date, etc.)
- › Cardinality – Describes how many instances of an element must be supplied.
- › 0 – Optional, can only supply 1.
- › 1 – Mandatory, must supply a single value.

## PBRF Staff Data file structure overview

The PBRF Staff Data file is submitted as a CSV file. The main reason for this is that staffing data can be easily extracted from Human Resource systems and exported to flat file structure formats like CSV. Staffing data was provided in the last Quality Evaluation round in this format and is easily uploaded and processed by the PBRF IT System.

The design of the PBRF Staff Data file structure and system processing assumes that the file:

---

<sup>1</sup> It should not be confused with the Ministry of Education's Annual TEO Staff Data Collection where staff data are submitted at the end of each calendar year or the collection of staffing data by Tribal for the New Zealand Benchmarking Tool.

- › is for a single TEO only
- › will contain at least one staff member but may contain many
- › will not contain a header row and will provide the fields in the order listed in the table below.

The table below summarises the basic structure of a PBRF Staff Data file.

Field Position	Field Name	Field Title
1	ProviderNumber	Provider Number
2	StaffID	Staff ID
3	NSN	National Student Number
4	DOB	Date of Birth
5	FirstName	First Name
6	PreferredFirstName	Preferred First Name
7	MiddleName	Middle Names
8	LastName	Family Name
9	PositionTitle	Position Title
10	FTE	Full-Time Equivalent
11	StartDate	Start Date
12	EndDate	End Date
13	NominatedAcademicUnit	Nominated Academic Unit
14	NewEmergResearcher	New and Emerging Researcher
15	PrevInstitID	Previous Provider ID
16	Gender	Gender
17	Ethnicity	Ethnicity

The following string is an example of how data may be supplied given the CSV format and field positions described above.

*Example:*

*7001,STFHV1,1234567890,19600115,Alex,Tony,Anthony Barry,Ballock,Head of Department,0.50,20140810,,School of Music,N,,M,211*

Notes:

- › The data are completely arbitrary and are only used to demonstrate an example of a CSV value in relation to the field names described above.
- › If data are entered into Excel and saved as a .CSV file, Excel 2010 will automatically insert double quotes around any field that contains a comma in the text.

### Transferred or concurrently employed staff

If a TEO has a PBRF-eligible staff member who has left the TEO between 15 June 2017 and 14 June 2018 (transferring staff member), or who is concurrently employed by another TEO and that TEO is submitting the EP (concurrently employed staff member), a record should be submitted in the PBRF Staff Data file. This will allow the TEC to match and proportion funding correctly.

The following assumptions will be made:

**PBRF-eligible staff who have transferred to another TEO will be defined by:**

- › no matching EP for this staff member from this TEO
- › there must be a matching EP (using NSN) from one other TEO
- › the inclusion of End Date implies the staff member transferred out on that date.

**Concurrently employed PBRF-eligible staff will be defined by:**

- › no matching EP for this staff member from this TEO
- › there must be a matching EP (using NSN) from one other TEO
- › End Date has no value
- › FTE must be less than 1.00.

An “orphan” record in the PBRF Staff Data file will result in a warning to the TEO. This condition will arise if there is no EP submitted from any TEO for the given NSN.

The PBRF IT System will determine if a staff member has transferred or is concurrently employed by applying the assumptions listed in above, and will indicate this on the TEO’s view of the staff data.

## PBRF Staff Data file specification details

Field	Description	Field Name	Cardinality	Data Type	Notes
Provider Number	Provider Number is a unique number that identifies a TEO. This is also known as the EDUMIS number.	ProviderNumber	1	number 4 characters	The field is used to identify TEOs. An empty value will result in an error and the record will not be processed. See <a href="#">Appendix A1</a> for a list of valid codes.
Staff Id	This is the identifier used by the TEO to reference the staff member. This may be an employee number or staff ID.	StaffID	0	string 10 characters	The field is used by the TEO to identify staff members. This field will be used as a reference field only. It may be useful in the event of a processing query for the TEC to be able to quote this when talking to a TEO.
National Student Number	A unique number issued by the National Student Index (NSI) that identifies a researcher who is shown as submitting an EP.	NSN	1	number 10 characters	The field is used by the PBRF IT System to identify staff members. Must be provided, as the NSN will be used as the primary mechanism to match researchers in the PBRF Staff Data file to those who have Evidence Portfolios. The TEC will use this number to search for the researcher in the National Student Index and check that the date of birth matches and that the NSI record found is a master record.
Date of Birth	The day, month, and year of birth of the staff member.	DOB	1	date 8 characters	The field is used in conjunction with the NSN to identify staff members. The field will be used in combination with the NSN to check that the NSN supplied matches the correct record on the National Student Index (NSI). If the Date of Birth in the staff data record does not match the date of birth in the NSI an error will result unless the NSI default date of birth has been used. In this case a warning will be given. Wherever possible actual date of birth should be

Field	Description	Field Name	Cardinality	Data Type	Notes
					<p>provided.</p> <p>Note that EP validation rules will check that the Date of Birth in the EP is the same as the Date of Birth in the Staff Data file.</p> <p>The field will be used by the TEC to report by age, e.g. staff in various subject areas by age.</p> <p>See <a href="#">Appendix A3</a> for valid format.</p>
First Name	The first or forename for the staff member.	FirstName	1	string 250 characters	<p>The field is used to identify staff members.</p> <p>An empty value will result in an error and the record will not be processed.</p>
Preferred First Name	The preferred first name for the staff member.	PreferredFirstName	0	string 250 characters	An optional field to aid in the identification of staff members known by their preferred first name.
Middle Name	This field consists of the staff member's middle or second forename(s).	MiddleName	0	string 250 characters	The field is used to identify staff members.
Family Name	This field consists of the staff member's surname (also known as last name or family name).	LastName	1	string 250 characters	The field is used to identify staff members.
Position Title	Job title of the staff member.	Title	1	string 250 characters	e.g. Senior Lecturer, Professor, Distinguished Professor.
Full-Time Equivalent	This field consists of the full-time equivalent (FTE) that the staff member is employed in the position at the TEO.	FTE	1	decimal 4 characters	<p>The FTE status for part-time staff is the percentage of full-time employment (to two decimal places).</p> <p>A full-time position or 1 FTE in the context of the Quality Evaluation is defined as 37.5 hours per week, which includes any non-research or teaching activities but excludes non-paid hours.</p> <p>The decimal point must be supplied e.g. 0.2 FTE</p>

Field	Description	Field Name	Cardinality	Data Type	Notes
					would be recorded as 0.20 and 1 FTE would be recorded as 1.00.  A staff member cannot be employed for more than 1.00 FTE.
Start Date	This field consists of the date on which employment of the staff member commenced at the TEO.	StartDate	1	date 8 characters	If the date of commencing employment (or contract) is in the last year this may affect the level of PBRF funding that can be obtained.  See <a href="#">Appendix A3</a> for valid format.
End Date	This field consists of the employment end date of the staff member at the TEO.	EndDate	0	date 8 characters	If a PBRF eligible staff member leaves within the year prior to the PBRF Staff Eligibility date (14 June 2018), the TEO may be able to obtain some benefit from the period of service that was conducted in the year.  In this situation, if an End Date is entered it must fall within the 12-month period prior to the PBRF Staff Eligibility date of 14 June 2018.  See <a href="#">Appendix A3</a> for valid format.
Nominated Academic Unit	This field consists of the employment position held by the PBRF eligible staff member at the TEO.	NominatedAcademicUnit	0	string 50 characters	This needs to be included if EndDate is not entered, otherwise optional.  This field will be used by the TEC for reporting of the 2018 PBRF Evidence Portfolio Peer Review.  Panel quality categories. This is one of the four levels of quality reporting that the TEC will provide both to TEOs and publicly. It will be used to aggregate the results of PBRF-eligible staff members within the same nominated academic unit at a TEO.  Actual choice of unit name is at the discretion of the TEO but should not contain the TEO name itself. Because it is used to aggregate results it is essential the TEOs understand the need for consistent use of the same name for any one Nominated Academic Unit – similar names or instances of wording reversals

Field	Description	Field Name	Cardinality	Data Type	Notes
					will be reported separately. Where a staff member is working jointly in more than one Nominated Academic Unit, the TEO should choose only one – the one where most time is spent – or which is most representative of the research undertaken.
New and Emerging Researcher	This field consists of a flag to identify staff who meet the criteria as a 'New and Emerging Researcher' for PBRF consideration.	NewEmergResearcher	0	string 1 character	This needs to be included if an EP is submitted for this researcher from this TEO, otherwise optional.  This flag denotes that a PBRF-eligible staff member meets the criteria as a 'New and Emerging Researcher' for the purposes of PBRF assessment.  See <a href="#">Appendix A2</a> for a list of valid codes.
Previous Provider ID	A four-digit code assigned by the Ministry of Education to uniquely identify a TEO which has merged with another TEO.	PrevInstitID	0	number 4 characters	The code is used for reporting purposes within the PBRF process, to identify significant organisational divisions within a TEO that have resulted from TEO merger.  See <a href="#">Appendix A1</a> for a list of valid codes.
Gender	The field is used to identify gender of staff members.	Gender	1	string 1 character	Any value other than M or F will default to U (Undeclared).  See <a href="#">Appendix A4</a> for a list of valid codes.
Ethnicity	The codes identify the ethnic group or groups to which a staff member belongs.	Ethnicity	0	string 3 characters	This needs to be included if an EP is submitted for this researcher from this TEO, otherwise optional.  From the ethnic groups specified, a single ethnicity is determined for each staff member by a system of priority recording based on the Statistics New Zealand standard classification of ethnicity.  See <a href="#">Appendix A5</a> for a list of valid codes.

## Appendix

### A1 TEO codes for SAC funded degree-delivering TEOs

TEO	Code
<b>Institutes of technology and polytechnics</b>	
Ara Institute of Canterbury	6006
Eastern Institute of Technology	6007
Manukau Institute of Technology	6010
Nelson Marlborough Institute of Technology	6011
Northland Polytechnic	6012
Open Polytechnic of New Zealand	6022
Otago Polytechnic	6013
Southern Institute of Technology	6015
Toi Ohomai Institute of Technology	6025
Unitec Institute of Technology	6004
Universal College of Learning	6009
Waikato Institute of Technology	6019
Wellington Institute of Technology	6008
Western Institute of Technology at Taranaki	6017
Whitireia New Zealand	6014
<b>Universities</b>	
Auckland University of Technology	7008
Lincoln University	7006
Massey University	7003
University of Auckland	7001
University of Canterbury	7005
University of Otago	7007
University of Waikato	7002
Victoria University of Wellington	7004
<b>Wānanga</b>	
Te Wānanga o Aotearoa	8630
Te Wānanga o Raukawa	9241
Te Whare Wānanga o Awanuiārangi	9386
<b>Private training establishments</b>	
Alphacrucis International College	8573
Anamata	9242
Auckland Institute of Studies	8530

TEO	Code
Bethlehem Tertiary Institute	8694
Carey Baptist College	8979
Good Shepherd College – Te Heparā Pai	8717
IPU New Zealand	8550
Laidlaw College Inc	8563
Media Design School	8192
New Zealand College of Chinese Medicine Limited	7282
New Zealand College of Chiropractic	8396
New Zealand School of Acupuncture and Traditional Chinese Medicine	9670
New Zealand Tertiary College	8619
Pacific International Hotel Management School	8457
Prema Charitable Trust	8341
SAE Institute	8174
South Pacific College of Natural Medicine	8642
Te Rito Maioha Early Childhood New Zealand	9520
Te Wānanga Takiura o Ngā Kura Kaupapa Māori o Aotearoa	8425
Toi Whakaari New Zealand Drama School	8502
Whitecliffe College of Arts and Design	8509

## A2 Simple Decision Codes

Description	Code
Yes	Y
No	N

## A3 Typical Date Format

The format of all dates requested will be: <b>CCYYMMDD</b> Example: 12 March 1965 will be formatted "19650312"	
Description	Code
Century Ranges from 19 to 21.	CC
Year ranges from 00 to 99. Please note that single digit <b>MUST</b> be prefixed with 0, e.g. 01 not just 1.	YY
Month ranges from 01 to 12. Please note that single digits <b>MUST</b> be pre-fixed with a 0, e.g. 01 not just 1.	MM
Day ranges from 01 to 31. Please note that single digits <b>MUST</b> be pre-fixed with a 0, e.g. 01 not just 1.	DD

#### A4 Gender Codes

Description	Code
Male	M
Female	F
Undeclared	U

#### A5 Ethnicity Codes

Description	Code
European nfd	100
New Zealand European	111
British and Irish	121
Dutch	122
Greek	123
Polish	124
South Slav	125
Italian	126
German	127
Australian	128
Other European	129
Māori	211
Pacific Peoples nfd	300
Samoan	311
Cook Islands Māori	321
Tongan	331
Niuean	341
Tokelauan	351
Fijian	361
Other Pacific Peoples	371
Asian nfd	400
Southeast Asian nfd	410
Filipino	411
Cambodian	412
Vietnamese	413
Other Southeast Asian	414
Chinese	421
Indian	431
Sri Lankan	441
Japanese	442

Description	Code
Korean	443
Other Asian	444
Middle Eastern	511
Latin American	521
African	531
Other Ethnicity	611
Don't Know	944
Refused to Answer	955
Repeated Value	966
Response Unidentifiable	977
Response Outside Scope	988
Not Stated	999